

UNIVERSIDADE FEDERAL DO PARANÁ

ORLANDO ALCANTARA SOARES

ESTRUTURA ARGUMENTAL:
UMA FUNDAMENTAÇÃO COMPUTACIONAL

CURITIBA
2010

ORLANDO ALCANTARA SOARES

ESTRUTURA ARGUMENTAL:
UMA FUNDAMENTAÇÃO COMPUTACIONAL

Tese apresentada ao Curso de Pós-Graduação em Letras,
Área de Concentração Estudos Linguísticos, Setor de
Ciências Humanas, Letras e Artes da Universidade Federal
do Paraná, como requisito parcial à obtenção do título
de Doutor em Linguística.

Orientador: Prof. Dr. José Borges Neto

CURITIBA
2010

Catálogo na publicação
Sirlei do Rocio Gdulla – CRB 9ª/985
Biblioteca de Ciências Humanas e Educação - UFPR

Soares, Orlando Alcantara
Estrutura argumental: uma fundamentação computacional /
Orlando Alcantara Soares. - Curitiba, 2010.
131 f.

Orientador: Prof. Dr. José Borges Neto
Tese (Doutorado em Letras) - Setor de Ciências Humanas,
Letras e Artes, Universidade Federal do Paraná.

1. Linguística. 2. Gramática gerativa – sintaxe. 3. Modelos
linguísticos. I. Título.

CDD 410

TERMO DE APROVAÇÃO

ORLANDO ALCANTARA SOARES

ESTRUTURA ARGUMENTAL: UMA FUNDAMENTAÇÃO COMPUTACIONAL

Tese aprovada como requisito parcial para a obtenção do título de Doutor em Letras, área de concentração Estudos Linguísticos, Setor de Ciências Humanas, Letras e Artes da Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientador: Prof. Dr. José Borges Neto
Universidade Federal do Paraná

Prof. Dr. Edson Emílio Scalabrin
Pontifícia Universidade Católica do Paraná

Prof. Dr. Luiz Arthur Pagani
Universidade Federal do Paraná

Prof.^a Dr.^a Maria Cristina Figueiredo Silva
Universidade Federal do Paraná

Prof. Dr. Maximiliano Guimarães
Universidade Federal do Paraná

Curitiba, 26 de outubro de 2010



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS HUMANAS, LETRAS E ARTES
COORDENAÇÃO DO CURSO DE PÓS GRADUAÇÃO EM LETRAS

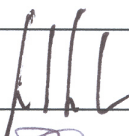

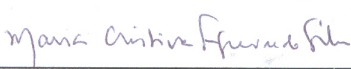

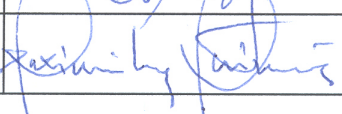
PARECER

Defesa de tese do doutorando ORLANDO ALCÂNTARA SOARES para obtenção do título de **Doutor em Letras**.

Os abaixo assinados JOSÉ BORGES NETO, LUIZ ARTHUR PAGANI, MARIA CRISTINA FIGUEIREDO SILVA, EDSON EMÍLIO SCALABRIN, e MAXIMILIANO GUIMARÃES arguíram, nesta data, o candidato, o qual apresentou a tese:

“ESTRUTURA ARGUMENTAL: UMA FUNDAMENTAÇÃO COMPUTACIONAL”

Procedida a arguição segundo o protocolo que foi aprovado pelo Colegiado do Curso, a Banca é de parecer que o candidato está apto ao título de **Doutor em Letras**, tendo merecido os conceitos abaixo:

Banca	Assinatura	APROVADO Não APROVADO
JOSÉ BORGES NETO		A
LUIZ ARTHUR PAGANI		Aprovado
MARIA CRISTINA FIGUEIREDO SILVA		Aprovado
EDSON EMÍLIO SCALABRIN		Aprovado
MAXIMILIANO GUIMARÃES		Aprovado

Curitiba, 26 de outubro de 2010


Prof. Dr. Maria José Foltran
Coordenadora

Dedicatória

Para a Gilda, com muito amor.

Agradecimentos

Meu tempo de preparação e pesquisa no doutorado foi marcado por alguns mestres a quem muito devo.

Primeiramente o meu orientador Borges (José Borges Neto), sem o qual esta tese não teria acontecido e com quem analisei as ideias aqui relatadas, em longas discussões. Foi o Borges quem me mostrou o ponto a investigar. Em uma de nossas conversas, a sua observação “*A grande contribuição do Chomsky à teoria da gramática foi a formalização do conceito de núcleo*” acendeu a lâmpada que iluminaria o resto da pesquisa. Aquela conversa de cantina se mostrou fundamental ao que viria depois.

A Mazé (Maria José Foltran) me introduziu aos estudos da Gramática Gerativa. Seu curso de Regência e Ligação mostrou uma quantidade de fenômenos completamente novos para mim e um modo de pensar com o qual eu ainda não tivera contato. Mesmo após a conclusão das disciplinas, a Mazé continuou a me abastecer de referências bibliográficas e a providenciar *papers*. Muito da literatura da pesquisa veio por seu intermédio.

O Max (Maximiliano Guimarães) teve um duplo papel. Sua grande erudição e enorme entusiasmo tornavam cada aula um dia de festa à qual eu comparecia com alegria. Suas aulas costumavam ser profundas discussões de temas definidos, sempre com literatura atualizada e de grande qualidade; outras vezes as indagações se tornavam altamente especulativas, abrindo horizontes; eram essas as aulas que eu mais gostava. O Max fez as perguntas que mais me estimularam: “*As variações linguísticas ocorrem sempre na posição de especificador. Por que?!?*”; “*Se temos o movimento, para que lambdas?!?*”. Bem, Max, para começar a dar uma resposta foi preciso escrever uma tese!

Muito me beneficiei das sugestões da Cristina (Maria Cristina Figueiredo Silva) a uma versão inicial do texto. Anotações suas me fizeram reescrever trechos importantes. A Cristina também me levou a evitar alguns excessos diacrônicos! Seu rigor conceitual corrigiu alguns “monólogos” demasiado informais (desculpe Cristina: demasiadamente informais) na minha escrita!

O texto da qualificação também se beneficiou da leitura atenta do Scalabrin (Edson Emílio Scalabrin), que deu sugestões não apenas para a parte computacional, sua especialidade. Sua fina capacidade de observação mostrou a necessidade de explicitar mais amplamente alguns conceitos e de reordenar alguns capítulos.

A todos o meu muito obrigado.

Resumo

Esta tese apresenta uma proposta de modelo computacional de representação linguística, aplicado à formação da estrutura argumental dos verbos e incidentalmente à dos nomes e adposições, considerando os adjetivos como uma categoria derivada, resultante de configurações estruturais. O modelo proposto formaliza a noção universal de núcleo, da qual núcleos específicos (como verbalizadores, nominalizadores e preposições) são derivados por parametrização. O núcleo é o formador de constituintes, estruturas portadoras de conteúdo semântico e fonológico. Grande parte da investigação analisa a semântica fundamental (mínima) de uma raiz e as condições de compatibilidade entre raízes e núcleos sintáticos, mostrando que, em consequência, determinadas estruturas são possíveis e outras impossíveis. Por seu elevado grau de abstração, o cálculo-lambda é o paradigma de representação da pesquisa: o seu pressuposto fundamental de não-distinção entre processo e dado o torna especialmente adequado à análise linguística. Duas consequências do modelo proposto são: (a) o processamento postergado da fonologia dos constituintes, sem ser necessário postular a inserção tardia da fonologia, como é proposto pela Morfologia Distribuída; (b) a não-necessidade de movimento de constituinte para a atribuição de caso nas estruturas verbais fundamentais.

Palavras-chaves: constituinte, derivação sintática, estrutura argumental, núcleo, representação linguística.

Abstract

This thesis presents the proposal of a computational model for linguistic representation, applied to the formation of the argument structure of verbs, and, incidentally, to that of nouns and adpositions, regarding the adjectives as a derived category, resulting from structural configurations. The proposed model formalizes the universal notion of nucleus (head), from which specific nuclei (such as verbalizers, nominalizers and adpositions) are derived by parameterizations. The nucleus is the constructor of constituents, structures bearing semantic and phonological contents. A large part of the investigation analyzes the fundamental (minimal) semantics of a root and the conditions of compatibility existent among roots and syntactic nuclei, showing that, as a consequence, certain structures are possible and others impossible. Due to its high degree of abstraction, lambda-calculus is the representational paradigm of the research: its fundamental presupposition of non-distinction between process and data makes it specially suitable to linguistic analysis. Two consequences of the proposed model are: (a) late processing of the constituents phonology, without being necessary to postulate the late insertion of phonology, as proposed by Distributed Morphology; (b) non-necessity of constituent movement for case assignment, in the domain of the fundamental verb structures.

Key-words: argument structure, constituent, head, linguistic representation, nucleus, syntactic derivation.

Lista de símbolos

Tipos Abstratos

Bool

S

$\langle S \rangle$

P

$\langle P \rangle$

Null

Significado

Tipo de valor booleano

Tipo de traço semântico

Tipo de uma sequência de traços semânticos

Tipo de traço fonológico

Tipo de uma sequência de traços fonológicos

Tipo do valor nulo

Operadores

Computacionais

β

error

φ

$\langle \varphi_i \rangle$

κ

null

str

σ

$\langle \sigma_i \rangle$

Significado

Função de seleção (função booleana)

Término anormal de uma derivação ou operação

Traço fonológico, operação fonológica

Sequência de traços fonológicos

Operador característico

Valor inexistente, valor vazio

Sequência de traços fonológicos

Traço semântico, operação semântica

Sequência de traços semânticos

Variáveis

Computacionais

b

e

h

i

k

p

s

Significado

Função de seleção

Argumento externo

Núcleo

Argumento interno

Operador característico

Sequência fonológica

Sequência semântica

Operadores Gramaticais

<i>ve</i>	Operador formador de elemento vocabular
<i>Ve</i>	Núcleo formador de elemento vocabular
<i>n</i>	Operador nominalizador
<i>N</i>	Núcleo nominalizador
<i>p</i>	Operador preposicional
<i>P</i>	Núcleo preposicional
<i>r</i>	Operador formador de raiz
<i>R</i>	Núcleo formador de raiz
<i>v</i>	Operador verbalizador
<i>V</i>	Núcleo verbalizador
<i>w</i>	Paradigma gramatical

Significado

Abreviaturas

DS	<i>Deep Structure</i> (Estrutura Profunda)
EPP	<i>Extended Projection Principle</i> (P. de Proj. Estendido)
Gen	Genitivo
GG	Gramática Gerativa
GU	Gramática Universal
LF	<i>Logical Form</i> (Forma Lógica)
MD	Morfologia Distribuída
MP	<i>Minimalist Program</i> (Programa Minimalista)
PB	Português-Brasileiro
PF	<i>Phonetic Form</i> (Forma Fonética)
Pl	Plural
PM	<i>Phrase Marker</i> (Marcador Frasal)
PP	<i>Projection Principle</i> (Princípio da Projeção)
Rel	Partícula relacional
Sg	Singular
SS	<i>Surface Structure</i> (Estrutura Superficial)
SVO	Sujeito-Verbo-Objeto

Significado

Sumário

I	Apresentação	1
1	Introdução	2
2	Programa da pesquisa	4
II	Fundamentos	7
3	Fundamentos computacionais	8
3.1	O cálculo-lambda	8
3.2	Elementos de cálculo-lambda	9
3.3	Operações morfológicas	13
3.4	Operadores com vários argumentos	14
3.5	Assinaturas funcionais de operadores	15
4	Ideias fundamentais	17
4.1	Um sistema computacional único	17
4.2	Teorias da formação argumental	22
4.2.1	Papeis semânticos	23
4.2.2	Estrutura de eventos	25
4.2.3	Hale e Keyser	28
4.2.4	Morfologia Distribuída	32
III	Noção formal de núcleo	33
5	Formalização da noção de núcleo	34
5.1	A noção de núcleo	34
5.2	Paradigma de representação	38
5.3	Formalização do núcleo universal	39
5.4	Núcleos e traços semânticos e fonológicos	47
5.5	Exemplo de núcleo parametrizado	48
5.6	O operador característico	50
5.6.1	Estrutura interna do operador característico	52
5.6.2	Exemplo de estrutura interna	53
5.7	O paradigma gramatical	55
5.8	Paradigma gramatical e argumentos do núcleo	56
5.9	Núcleo e diagramas	57
5.10	Núcleo e categorias gramaticais	58

5.11	Definição formal com assinaturas funcionais	59
5.12	Propriedades estruturais do contexto	60
IV	Derivações sintáticas	64
6	Estruturas básicas	65
6.1	Tipos de núcleos	65
6.2	Relacionamento funcional entre os tipos de núcleos	66
6.3	O elemento vocabular	68
6.3.1	Evidência empírica	69
6.3.2	Evidência do semita	69
6.3.3	Evidência do hebraico	70
6.4	Resumo provisório	74
6.5	Derivação lexical básica	76
6.5.1	Derivação do elemento vocabular	77
6.5.2	Núcleos e traços semânticos	79
6.5.3	Derivação da raiz	80
6.5.4	Derivação do nome	83
6.5.5	Derivação genérica de nome	86
7	Estruturas estativas	88
7.1	Raízes estativas básicas	88
7.2	Derivação estativa básica	89
7.3	Derivação nominal	93
7.4	Verbalização estativa	95
8	Estruturas ativas	97
8.1	Estrutura inacusativa	97
8.2	Noção de voz ativa	99
8.3	Derivação verbal inacusativa	100
8.3.1	O sujeito inacusativo	102
8.3.2	Estatividade e particípio	104
8.4	Verbos inergativos	105
8.5	Verbos transitivos	108
8.6	Verbos alternantes causativo-incoativos	112
8.6.1	Alternância de verbos ativos	112
8.6.2	Alternância de verbos estativos	118
8.6.3	Alternância, transitividade e incoatividade	125
8.7	Verbos ditransitivos	127
8.7.1	Os verbos biargumentais	128
8.7.2	A sugestão de Chomsky	128
8.7.3	A proposta de Larson	129
8.7.4	A proposta de Hale e Keyser	130
8.7.5	A proposta de Harley	134
8.8	Preposições	136
8.8.1	A estrutura verbal biargumental	136
8.8.2	Verbos gramaticais biargumentais	137
8.8.3	Estrutura de duplo-complemento	137

8.8.4	Estrutura de duplo-objeto	139
8.8.5	Comparação com a solução de Harley	142
8.8.6	Os verbos denominais	143
8.8.7	A preposição locativa ativa	144
8.8.8	Derivação da sentença média	146
8.8.9	O verbo gramatical <i>pôr</i>	147
8.8.10	Derivação da sentença transitiva	148
8.8.11	A preposição condicional	149
8.8.12	Preposição estativa básica	151
8.9	Voz	153
8.9.1	Voz ativa	155
8.9.2	Vozes média e passiva	156
8.9.3	O sujeito dos verbos biargumentais	160
8.9.4	A semântica da voz média	162
9	O Princípio de Projeção Estendido	164
9.1	EPP e processos	165
9.2	EPP e estatividade	168
9.3	Modelo proposto e EPP	170
V	Conclusão	171
10	Considerações finais	172

Parte I

Apresentação

Capítulo 1

Introdução

O objetivo deste trabalho é a proposta de um modelo computacional de representação linguística, aplicado à formação da estrutura argumental dos verbos e, incidentalmente, à dos nomes e preposições.¹

Será investigada a semântica fundamental (mínima, em sentido a ser definido) de uma raiz e as condições de compatibilidade da raiz com os núcleos (operadores) gramaticais, para a formação da estrutura argumental, procurando mostrar por que determinadas estruturas são possíveis e outras impossíveis.

O modelo computacional baseia-se nas noções formais de núcleo e de constituinte. A partir destas noções são derivadas as estruturas verbais fundamentais (*inacusativas, inergativas, transitivas, ditransitivas, estativas e existenciais*)² e o tema correlato da alternância de transitividade *causativa-incoativa*. Um dos pontos fundamentais é a análise do importante papel que as *preposições* e a *voz* (*ativa, média e passiva*) desempenham no processo de formação das estruturas verbais.

Além de propor um modelo computacional, a pesquisa encara a própria linguagem como um fenômeno computacional. O modelo proposto deve ser compreendido como uma simulação do fenômeno da linguagem, uma máquina conceitual que busca reproduzir, abstratamente, o fenômeno linguístico, como supostamente ocorre na mente de um sujeito falante.

Considerar a linguagem como um fenômeno computacional implica que ela não pode ser incompatível com as leis computacionais fundamentais. Neste sentido, realizamos uma investigação dos fenômenos linguísticos a partir de um paradigma mínimo de representação que somente emprega: (a) a ideia computacional mais elementar, a *interação entre dois termos*; (b) o princípio mais fundamental da construção de uma teoria, a *abstração*, para representar os objetos que interagem; e (c) a noção de *lei*, que rege a interação entre os termos.

Este paradigma mínimo está filosoficamente fundamentado na semiótica de Peirce

¹Segundo o modelo proposto, somente nomes e verbos constituem categorias lexicais: (a) os adjetivos são estruturalmente caracterizados, formados por uma relação sintática especial entre constituintes de determinado tipo (basicamente raízes) e constituintes nominais; (b) as proposições pertencem ao sistema funcional da língua.

²As estruturas estativas e as existenciais são consideradas como categorias próprias.

(Peirce (1867), Santaella (2000)) e no paradigma computacional de Church (Church (1941), Barendregt (1997)), o cálculo-lambda. O cálculo-lambda realiza o princípio de minimalidade computacional, em que cada expressão é simultaneamente uma abstração e uma interação.

A pesquisa adota um viés gerativo chomskiano, segundo o espírito da proposta minimalista de Chomsky (1993a).

Tanto a semiótica de Peirce quanto o cálculo-lambda são neutros com relação aos campos a que se aplicam e, por seu minimalismo epistemológico, mostram-se especialmente adequados a fundamentar uma representação de caráter minimalista do fenômeno da linguagem.

Alguns resultados já estabelecidos pela teoria gerativa serão tomados como ponto de partida. Por exemplo, não haverá a preocupação de provar que o sujeito de um verbo inergativo tem origem no argumento externo do verbo, ou que, em uma sentença inacusativa, o sujeito provém do argumento verbal interno. Tais fatos são considerados como suficientemente fundamentados, não necessitando de argumentação suplementar. As análises buscarão formalizar computacionalmente tais estruturas e investigar as consequências do modelo proposto para a teoria da formação da estrutura argumental.

O foco do trabalho são os fenômenos de formação argumental, o que deixa de fora temas relacionados, como os mecanismos de atribuição de caso e de concordância. Haverá uma análise da caracterização estrutural do caso, entretanto, embora sendo de grande importância para a teoria da gramática, os mecanismos da atribuição morfológica de caso ultrapassam os limites deste relato.

Para a maioria dos exemplos utiliza-se o português-brasileiro (PB), no que se pode razoavelmente denominar “dialeto-padrão”, a língua espontaneamente falada em ambiente doméstico pela maioria das pessoas de escolaridade média do Brasil, ainda que de modo não homogêneo; eventualmente, são também utilizadas outras variedades dialetais do PB, assim como outros idiomas.

Embora as estruturas analisadas sejam majoritariamente ilustradas por expressões do PB, o modelo proposto pretende ter validade universal, procurando simular uma parte do sistema computacional da Gramática Universal (GU), compreendida essencialmente no seu sentido chomskiano.

Capítulo 2

Programa da pesquisa

Neste trabalho proponho uma abordagem computacional de representação linguística em que:

- (a) A formação lexical e a derivação sentencial ocorrem exclusivamente na sintaxe, em um sistema computacional único;
- (b) As operações sintáticas resultam da interação entre *núcleos* (operadores pertencentes ao sistema funcional da língua) e *constituintes* (estruturas portadoras de informação semântica e fonológica);¹
- (c) As configurações estruturais dos constituintes determinam a interpretação temática de seus argumentos.

A abordagem segue um programa de pesquisa de amplo alcance proposto por Hale & Keyser (1993) e ampliado pela Morfologia Distribuída (Halle & Marantz (1993), Marantz (1996), Marantz (1998), Harley & Noyer (Abril 1999)), em seu projeto de explorar as consequências do pressuposto minimalista de Chomsky (1995) de que a faculdade da linguagem possui um único sistema computacional, e, em consequência, todas as operações gramaticais são processadas na sintaxe.

A hipótese de um único sistema computacional para a formação lexical e sentencial é um dos princípios fundamentais de Chomsky (1995), e, por suas implicações, é provavelmente incompatível com a existência de um subsistema computacional específico para formação do léxico, no sentido proposto pela *Hipótese Lexicalista*, de *Remarks on Nominalization* de Chomsky (1970), e retomado pela Teoria da Regência e Ligação, de Chomsky (1981). Na presente proposta, o mesmo sistema computacional deriva as estruturas lexicais e sentenciais.

Toda a presente pesquisa parte do princípio gerativista de que a informação (funcional, semântica e fonológica) pertencente a um constituinte determina o seu comportamento sintático, isto é, determina as estruturas que podem ser geradas a partir do constituinte.

Este princípio pode ser compreendido de diferentes modos.

¹Os conceitos de núcleo e constituinte serão desenvolvidos ao longo do texto.

(a) Uma ideia de longa tradição no gerativismo propõe a associação de um item lexical (inicialmente, apenas verbos) a uma lista de especificações temáticas (como *Agente*, *Tema*, *Meta*), como nos *case frames* de Fillmore (1968), e, posteriormente, nas *grades temáticas* de Stowell (1981), e na *Teoria da Regência e Ligação* de Chomsky (1981).

(b) Teorias semânticas alternativas, como a de Levin & Rappaport Hovav (2005), propõem substituir as grades temáticas por estruturas mais complexas, que descrevem os eventos indicados pelos itens lexicais (basicamente verbos).

As duas abordagens acima associam os argumentos das estruturas semânticas lexicais (respectivamente, grades temáticas e descrições de eventos) aos argumentos das estruturas sintáticas, por meio de uma *teoria de associação (linking theory)*. Ou seja, para resolver o problema da estrutura argumental dos itens lexicais, desenvolvem três teorias: uma teoria semântica lexical; outra, sintática; e uma terceira, de associação entre a teoria semântica e a sintática. Em alguns casos, como em Levin & Rappaport Hovav (2005), a teoria semântica torna-se tão ou mais complexa quanto a teoria sintática.

(c) A proposta de Hale & Keyser (1998) dispensa uma teoria de associação sintático-semântica e simplifica a semântica lexical: os itens lexicais diretamente especificam as estruturas sintáticas em que podem ser inseridos. Ao combinar, sintaticamente, operadores gramaticais (como V, N e P) e raízes, as estruturas sintáticas são geradas diretamente a partir de informação lexical, sem a interveniência de uma associação sintático-semântica.

A teoria de Hale e Keyser (H&K) é um passo adiante em direção à adoção de um sistema computacional único, embora os dois autores proponham dois subsistemas computacionais, um para a derivação lexical e outro para a derivação sentencial. Entretanto, no meu entender, a abordagem de H&K apresenta problemas para lidar com verbos transitivos lexicais (aqueles em que a transitividade deriva de propriedades da raiz).²

Há dificuldades estruturais para modelar, por um lado, verbos denominais como *laugh* (Hale & Keyser 2002, p. 15), e deadjetivais (alternantes) como *clear*, p. 16 ('The screen cleared', 'I cleared the screen'), e, por outro lado, verbos transitivos como *cut*, p. 44 ('I cut the bread'). As dificuldades resultam da criação automática de um sujeito (externo) para *clear* (para provocar a alternância de transitividade) e da necessidade de impedir que o mesmo sujeito seja criado para *cut* (para evitar que a alternância ocorra). Tal fato leva a que verbos transitivos, como *cut*, sejam representados com um sujeito interno (gerado a partir de um argumento introduzido na derivação abaixo do operador verbalizador) — talvez seja a única abordagem, na teoria linguística, em que o sujeito de um verbo transitivo lexical tenha origem abaixo do núcleo verbal. Os dois autores ligam tal fenômeno à ocorrência paralela de formas sentenciais transitivas e médias ('I cut the bread', 'This bread cuts easily').

²Como *lavar*, em que uma raiz ativa requer um complemento, diferentemente de *engavetar*, em que a raiz é alheia a fenômenos de complementação e transitividade. As diferenças estruturais entre verbos como *lavar* e *engavetar* constituem um dos temas do trabalho.

No modelo que proponho, do mesmo modo que para H&K, o conteúdo semântico lexical, combinado com núcleos (operadores gramaticais), também dá origem diretamente às estruturas sintáticas, dispensando uma teoria de associação entre a semântica lexical e a estrutura sintática resultante. Mas, no modelo proposto, um constituinte não especifica as configurações sintáticas em que pode ser inserido, como é postulado por H&K: o constituinte é completamente agnóstico quanto aos ambientes de que pode participar. Minha solução é isenta de problemas combinatórios como o apontado no parágrafo anterior, como pretendo mostrar. A semântica lexical disponibiliza informação para o processamento de núcleos, mas não determina os núcleos que a podem processar. Isto significa que o processamento dos traços semânticos e fonológicos é puramente interpretativo; um determinado núcleo pode até mesmo violar o que seria o “comportamento esperado” de um dado constituinte, em função da semântica ou da fonologia do constituinte. Este é o espaço, por exemplo, da variação linguística em uma língua específica.

A semântica lexical de um constituinte, no modelo proposto, compõe-se da combinação de quatro traços semânticos que expressam: (a) a *estatividade* ou *atividade* do constituinte, por meio de um dos traços *st* ou *ac*; (b) a possibilidade de o constituinte ser compatível com um complemento, indicada pelo traço semântico de *incidência*, *in*; (c) o tipo de causalidade associada ao constituinte, *interna* ou *externa*, indicada pelos traços *ic* e *ec*.³ A fundamentação, o significado e o uso dos traços semânticos, como também as derivações que eles possibilitam, são tema de todo o trabalho.

É uma questão empírica determinar se os traços semânticos acima mencionados são suficiente para especificar a derivação sintática das estruturas argumentais. Se as análises desenvolvidas no restante do texto estiverem corretas, eles são suficiente para derivar as construções verbais fundamentais: (a) as estativas e existenciais; e (b) as ativas (inergativas, inacusativas, transitivas lexicais e gramaticais, causativas lexicais alternantes, ditransitivas, denominais).⁴

Ao adotar um sistema sintático único, a presente proposta é semelhante à da *Morfologia Distribuída* (MD); entretanto, dela difere fundamentalmente por não recorrer ao princípio de *late insertion* (*inserção tardia* de material fonológico, Halle & Marantz (1994, p. 275)). Na MD, a inserção de expressões fonológicas lexicais ocorre após as operações sintáticas, no processo denominado *Spell-Out*. No nosso caso, o conteúdo fonológico de um item lexical é inserido no primeiro passo de formação do item; o sistema computacional assegura o processamento postergado (*late processing*) do material fonológico lexical e gramatical, como consequência da estrutura dos constituintes, tornando desnecessária a hipótese de *late insertion*. Não há, assim, a necessidade de um algoritmo pós-sintático de inserção de expressões fonológicas.

Um resultado interessante, também consequência do modelo computacional, é que não há a necessidade de movimento de constituintes para a atribuição de caso nas estruturas verbais fundamentais.

³Os traços *ec* e *ic* são independentes, configurando-se assim um conjunto de quatro traços semânticos de um item lexical: *ac* ou *st*, *in*, *ec*, *ic*.

⁴Algumas designações, como *estruturas transitivas lexicais e gramaticais* e *estruturas causativas lexicais*, serão explicadas em capítulos vindouros.

Parte II

Fundamentos

Capítulo 3

Fundamentos computacionais

3.1 O cálculo-lambda

O ano de 1936 viu surgir dois modelos computacionais, o *cálculo-lambda*, de Alonzo Church, e as *máquinas de Turing*, de Alan Turing. Em 1937, Turing demonstrou que os dois modelos são equivalentes (Turing 1937): toda função computável por um dos modelos também é computável pelo outro. Atualmente se aceita que ambos os modelos definem a mesma classe de funções, as funções computáveis por algoritmos; neste sentido, ambos os modelos são *modelos universais* de computação.

As máquinas de Turing implementam o conceito de dispositivo dotado de memória, programa armazenado e dado, realizando abstratamente a ideia de processo executado por um equipamento. Neste sentido, grande parte dos computadores atuais são modelos físicos de máquinas de Turing, embora sejam comumente denominados *máquinas de von Neumann*, em homenagem a John von Neumann, que definiu a arquitetura desta classe de computadores a partir das ideias de Turing.

O cálculo-lambda formaliza a noção de computação abstrata, independente de dispositivo físico. Sua ideia fundamental é a de que um termo (um algoritmo) se aplica a outro termo (um valor), produzindo um resultado, não havendo entretanto diferença essencial entre o que é algoritmo e o que é valor.¹

Por seu grau de abstração, o cálculo-lambda é o paradigma computacional da presente pesquisa. O pressuposto fundamental de não-distinção entre processo e dado o torna especialmente adequado à análise linguística: o mesmo termo linguístico pode atuar como dado ou como processo, dependendo do contexto em que figura. A independência de dispositivos possibilita ao cálculo-lambda representar a estrutura abstrata do fenômeno linguístico, desvinculada de hipóteses a respeito do meio em que essa estrutura se realiza.

¹Os parágrafos precedentes têm a finalidade de situar o cálculo-lambda como modelo universal de computação (a ideia de algoritmo independente de dispositivo de implementação), contrastando-o com outro modelo universal de computação, as máquinas de Turing (a ideia abstrata de uma máquina programável dotada de memória). Para uma introdução às máquinas de Turing, remetemos o leitor a Carnielli & Epstein (2006, p. 105). Uma comparação técnica entre o cálculo-lambda e as máquinas de Turing extrapola o objetivo desta pesquisa.

Na próxima seção são apresentados os elementos do cálculo-lambda utilizado no restante do trabalho, focalizando as noções fundamentais e a aplicação computacional. Na medida do possível, evitou-se a formalização matemática mais técnica, a qual, embora necessária à fundamentação teórica, não apresenta interesse imediato a aplicadores da teoria. Os leitores interessados no aprofundamento do assunto podem consultar Partee et al. (1990) (para uma apresentação dirigida a linguistas), Barendregt (1984) (como uma referência enciclopédica do tema) e Barendregt (1990) (para um tratamento teórico resumido).

3.2 Elementos de cálculo-lambda

A sintaxe do cálculo-lambda é notavelmente simples, baseada em apenas duas operações, *aplicação* e *abstração*. A *aplicação* de uma expressão f a uma expressão a é denotada pela expressão

$$f a$$

em que se interpreta f como uma função (um algoritmo, um processo, um operador) que processa o dado (valor) a .

Para exemplificar, consideremos a *sequência de caracteres* (*string*)² ‘menin’ e o processo que designaremos por *fem*, que concatena o caráter ‘a’ à direita de uma *string*. Se aplicarmos o processo *fem* à sequência ‘menin’, obteremos a aplicação

$$fem\text{ ‘menin’}$$

que produz como resultado a sequência de caracteres

$$\text{‘menina’}.$$

A produção de um resultado por uma aplicação funcional será indicada pelo símbolo \mapsto . Assim, a expressão

$$fem\text{ ‘menin’} \mapsto \text{‘menina’}$$

mostra um termo (‘menina’) como resultado de uma aplicação funcional ($fem\text{ ‘menin’}$). A aplicação de uma expressão-lambda a outra é também uma expressão-lambda. No exemplo dado, a aplicação $fem\text{ ‘menin’}$ é, ela própria, uma expressão-lambda. O processo em que uma aplicação funcional dá origem a um resultado é chamado de “redução”. Diz-se assim que a expressão $fem\text{ ‘menin’}$ foi “reduzida” para a expressão ‘menina’. Como veremos neste trabalho, a derivação (sintática) de uma expressão linguística é uma sequência de reduções.

É possível compreender a aplicação funcional $fem\text{ ‘menin’}$ de dois modos: como a pró-

²Utilizaremos indistintamente, como sinônimos, os termos *sequência de caracteres* e *string*; um caráter particular, como ‘a’, é também uma *string* e, ao mesmo tempo, uma sequência (unitária) de caracteres. Postularemos que existem *strings* “vazias”, aquelas que não têm caracteres, em analogia com o conceito de *conjunto vazio*.

pria expressão *fem* ‘menin’ ou como o seu resultado ‘menina’; cada expressão é considerada como denotadora do mesmo fenômeno, observado em diferentes níveis computacionais.

A segunda operação do λ -cálculo,³ a *abstração*, é efetivamente um esquema de operações. A *abstração*

$$\lambda x.f x$$

é uma expressão em que uma aplicação $f x$ é “dependente de x ”, um valor não determinado. Para cada valor de x tem-se uma aplicação diferente.

Por exemplo, a abstração

$$\lambda x.fem x$$

denota o processo de sufixação de ‘a’ a uma *string* x não definida; isto é, denota genericamente as aplicações que possuem a forma $fem x$.

Ao se aplicar uma abstração $\lambda x.f x$ a um valor a , tem-se a aplicação

$$(\lambda x.f x) a$$

que se reduz para $f a$, também esta uma aplicação. Na expressão acima, os parênteses são necessários para delimitar a abstração.

Na aplicação $(\lambda x.f x) a$ ocorre a *substituição* de x por a , resultando $f a$:

$$(\lambda x.f x) a = f a.$$

Exemplificando: a sufixação de ‘a’ a ‘menin’ pode ser representada por meio de uma combinação de abstração e aplicação:

$$(\lambda x.fem x) 'menin' = fem 'menin'.$$

Qualquer um dos dois termos da igualdade acima produz o resultado ‘menina’.

Na abstração $\lambda x.f x$ diremos que f e x são *variáveis*: x é uma variável *ligada*, por estar “sob efeito” do λ , e f é uma variável *livre*, por não figurar previamente no λ .

As variáveis ligadas apresentam uma importante propriedade: a expressão $\lambda x.f x$ e $\lambda w.f w$ são equivalentes, e ambas são equivalentes a $\lambda y.f y$. Não é relevante o nome dado às variáveis ligadas (no caso, x , w ou y), a estrutura das expressões é exatamente a mesma; a estrutura de uma expressão-lambda independe do nome dado às variáveis.

Uma abstração pode conter outra abstração. É possível uma expressão como

$$\lambda y.(\lambda x.(f x y))$$

em que os parênteses têm a função de delimitar o escopo dos operadores lambda. Existem assim duas abstrações nessa expressão, uma para cada λ .

³Utilizaremos indistintamente as formas λ -cálculo ou cálculo-lambda.

Utilizemos as abstrações acima para representar o modo como atua a *prefixação*, a operação que concatena uma *string* à esquerda de outra *string*. Tal operação será denominada *prefix*. Assim, de acordo com a conceituação desenvolvida até o momento, tem-se:

$$\text{prefix} \text{'re' 'marca'} \mapsto \text{'remarca'},$$

em que a operação *prefix* se aplica a dois valores, 're' e 'marca', para produzir o resultado 'remarca'. Como operação que se aplica a dois valores, *prefix* pode ser representada por meio de uma abstração com duas variáveis ligadas:

$$\lambda y.(\lambda x.(\text{prefix}xy)).$$

Podemos agora fazer uma prefixação em duas etapas, uma para cada variável da expressão:

$$(\lambda y.(\lambda x.(\text{prefix}xy))) \text{'marca'}.$$

Na expressão acima, o valor 'marca' substitui a variável *y*; com a substituição o λy é eliminado da expressão, resultando então:

$$\lambda x.(\text{prefix}x \text{'marca'}),$$

uma expressão-lambda “em *x*”. A variável *x* será agora substituída pela *string* 're':

$$(\lambda x.(\text{prefix}x \text{'marca'})) \text{'re'}.$$

Ao se substituir a variável pelo valor, tem-se a eliminação do segundo λ , com o resultado

$$\text{prefix} \text{'re' 'marca'},$$

que se reduz para 'remarca', de acordo com a definição do operador *prefix*.⁴

Uma expressão com lambdas “embutidos”, como $\lambda y.(\lambda x.(\text{prefix}xy))$, pode ser escrita de um modo simplificado, porém equivalente por convenção, com apenas um operador λ , como é mostrado abaixo:

$$\lambda yx.\text{prefix}xy.$$

Em uma expressão-lambda com várias variáveis é comum os valores das variáveis serem repassados à expressão de uma única vez, como se mostra a seguir. É preciso ter em conta, entretanto, que os valores são substituídos passo-a-passo, em etapas, segundo o processo anteriormente analisado.

$$(\lambda yx.\text{prefix}xy) \text{'marca' 're'}.$$

⁴Operações como *prefix* e *fem* são utilizadas neste capítulo com finalidade ilustrativa de aplicação do lambda-cálculo. Mas algumas nuances podem ser comentadas: (a) *prefix* é compreendida como uma operação morfológica, daí o fato de aplicar-se binariamente a dois termos por ela reunidos; (b) *fem* é supostamente uma operação semântica, modalizadora (classificadora) de um termo, por isso o seu caráter monoargumental.

Além dos conceitos já definidos (aplicação, abstração, variável, valor), algumas outras definições são importantes para o restante do texto. Considerando a expressão-lambda acima:

- (a) As variáveis ligadas, x e y , são também denominadas *argumentos* da expressão, de acordo com uma tradição que remonta a Frege;⁵
- (b) A expressão λyx será denominada uma *parametrização* da expressão-lambda (uma convenção particular desta pesquisa);
- (c) x e y , como figurantes de λyx , serão ditos *parâmetros* da expressão (também uma convenção particular da pesquisa);
- (d) Por analogia com o vocabulário da química, x e y , figurantes de λyx , serão também eventualmente denominados *valências* da expressão (um modo comum de expressar as posições a serem preenchidas em um predicado, usual em lógica e em algumas vertentes da linguística, como a *Gramática de Valências*; também faz parte do vocabulário informal de linguistas pertencentes a outras filiações teóricas);
- (e) Valores como 'marca' e 're' serão chamados de *valores de argumentos* da expressão, uma designação habitual em λ -cálculo; como há uma correspondência um-a-um entre valores de argumentos e argumentos de uma expressão-lambda, quando não houver ambiguidade, os valores de argumentos serão chamados de argumentos da expressão (um conveniente abuso de linguagem, frequentemente utilizado por este autor).

Como se pode observar, há uma variedade de termos referentes às variáveis ligadas; a escolha do termo a utilizar em geral é feita de acordo com o contexto da análise, em conformidade com certas tradições e para realçar determinados sentidos. As definições tomaram por base um exemplo específico, mas são completamente gerais, sendo válidas, *mutatis mutandis*, para expressões-lambda com qualquer número e nomes de variáveis, como as expressões que possuem a parametrização $\lambda bkspie$, característica da noção de núcleo a ser formalizada alguns capítulos adiante.

Voltemos agora ao operador *fem* e definamos um outro operador, *plural*, que sufixa o caráter 's' a uma *string*. Podemos então obter a sequência 'meninas' pela aplicação ordenada dos dois operadores à sequência 'menin', do seguinte modo:

$$plural(fem\text{'menin'}).$$

A aplicação de um operador (como *plural*) ao resultado de outro operador (como *fem* 'menin', que reduz para 'menina') é uma importante técnica do cálculo-lambda. A aplicação sucessiva de operadores pode ser redefinida em termos da *composição de operadores* (ou de funções), conceito denotado pelo operador \circ . Utilizando-se a composição de operadores, a aplicação

⁵O operador *prefix* é uma variável livre cujo valor (o processo de prefixação) é dado pelo ambiente em que a expressão-lambda está inserida.

funcional acima pode ser expressa de modo equivalente por

$$(plural \circ fem) 'menina'.$$

O conceito de composição de operadores é expresso de modo preciso por meio da definição:

$$(g \circ f)x = g(fx).$$

A composição de operadores é um dos recursos mais fundamentais utilizados na análise linguística da presente pesquisa.

3.3 Operações morfológicas

O cálculo-lambda será agora usado para definir as operações morfológicas de *sufixação*, *prefixação* e *infixação*. Essas operações podem ser entendidas em termos de uma operação fonológica mais básica, a *concatenação*, que denotaremos pelo símbolo \parallel : a concatenação do termo ‘a’ ao termo ‘b’ será formalizada pela expressão $\parallel 'a' 'b'$, que dá como resultado ‘ab’. A operação \parallel é definida pela expressão-lambda

Definição 1. $\parallel = \lambda xy.xy$.

Utilizando a operação \parallel , a sufixação da sequência fonológica ‘a’ à sequência fonológica ‘menin’ é especificada por meio da operação *suffix*, como a seguir

$$suffix 'a' 'menin' = \parallel 'menin' 'a' = 'menina'.$$

Considerando o exemplo acima, a operação *suffix* recebe a seguinte definição em cálculo-lambda.

Definição 2. $suffix = \lambda xy.\parallel yx$.

Assim, ‘menina’ tem a derivação

$$\begin{aligned} suffix 'a' 'menin' &= (\lambda xy.\parallel yx) 'a' 'menin' \\ (\lambda xy.\parallel yx) 'a' 'menin' &\mapsto 'menina' \end{aligned}$$

em que (a) a primeira linha da derivação expressa a equivalência entre o operador *suffix* e a sua definição por meio da expressão-lambda $\lambda xy.\parallel yx$ e (b) a segunda linha mostra a redução para ‘menina’ obtida pela aplicação da expressão-lambda aos seus dois argumentos.

A relação que um operador (como *suffix*) tem com sua definição (como $\lambda xy.\parallel yx$) é de *sinonímia*, fato ilustrado pela primeira linha da derivação: se um operador é usado em um contexto computacional (contexto de aplicação funcional), sua definição igualmente pode ser

usada, e vice-versa.

Pode-se definir a operação de prefixação *prefix*, já encontrada anteriormente, de modo similar ao feito para a definição de *suffix*.

Definição 3. $prefix = \lambda xy. || xy$.

Por esta definição, o termo ‘remarca’ é derivado como a seguir:

$$\begin{aligned} prefix \text{ ‘re’ ‘marca’ } &= (\lambda xy. || xy) \text{ ‘re’ ‘marca’ } \\ (\lambda xy. || xy) \text{ ‘re’ ‘marca’ } &\mapsto \text{ ‘remarca’ }. \end{aligned}$$

A operação de infixação, embora computacionalmente similar às duas anteriores, é de definição mais complexa, uma vez que é necessário especificar a primeira e a segunda partes do termo que receberá o infixo. Por exemplo, consideremos a derivação da sequência fonológica ‘men’, a partir de uma hipotética sequência mais básica ‘mn’ mediante a infixação fonológica do ‘e’.⁶ A derivação de ‘men’ é então

$$infix \text{ ‘e’ ‘mn’ } \mapsto \text{ ‘men’ }.$$

Nesta derivação, a sequência ‘mn’ possui uma primeira parte, ‘m’, que é determinada pela operação *first*, em que $first \text{ ‘mn’ } = \text{ ‘m’ }$. Analogamente, a segunda parte da sequência é determinada pela operação *second*, tendo-se assim $second \text{ ‘mn’ } = \text{ ‘n’ }$. Levando-se em conta a decomposição em duas partes do segundo argumento da operação *infix*, a definição em cálculo lambda resulta

Definição 4. $infix = \lambda xy. prefix (suffix x (first y)) (second y)$.

Com a definição acima,

$$infix \text{ ‘e’ ‘mn’ } \mapsto \text{ ‘men’ }.$$

Outras operações morfológicas, como as alternâncias vocálicas e consonantais, podem ser também definidas em cálculo-lambda; tais definições serão deixadas à parte, por se situarem fora do objetivo imediato da pesquisa.

3.4 Operadores com vários argumentos

Munidos das definições da seção anterior, podemos definir a operação *fem* que sufixa a *string* ‘a’ a uma sequência fonológica:

$$fem = suffix \text{ ‘a’ }.$$

A operação *fem*, ao ser definida por meio de *suffix*, é definida, em última análise, a partir de $||$,

⁶A operação de infixação apenas marginalmente pertence ao sistema morfológico do português, se é que pertence, mas constitui um processo produtivo na família de línguas semíticas. A definição da operação *infix*, especificada acima, como já dito anteriormente, tem finalidade meramente ilustrativa; extrapola os limites deste trabalho um tratamento rigoroso das operações morfológicas.

a concatenação de *strings*. Assim: *fem* uma operação de um argumento, é definida a partir de *suffix*, uma operação de dois argumentos, em que o valor do primeiro argumento é dado (igual a 'a').

Temos então uma propriedade fundamental das operações: uma operação de dois argumentos, como *suffix*, ao se aplicar a um valor, como 'a', continua a ser uma operação, reduzida porém a um único argumento. O fato é mostrado a seguir.

$$\begin{aligned} \text{suffix 'a'} &= (\lambda xy. \|yx) \text{'a'} \\ &= \lambda y. \|y \text{'a'} \end{aligned}$$

Isto é:

$$\text{fem} = \text{suffix 'a'} = \lambda y. \|y \text{'a'},$$

em que *fem* é uma operação com uma única valência aberta, a da variável ligada *y*, uma vez que a variável *x* foi preenchida por 'a'.

Em termos gerais, em uma aplicação *fxy*, *f* é uma função que se aplica a *x*, e *fx* é uma função que se aplica a *y*. Ou seja: *f*, uma função de dois argumentos, pode ser conceituada como a aplicação sucessiva de duas funções monoargumentais. O fenômeno é universal, válido para funções com qualquer número de argumentos.

A ideia de que funções com mais de um argumento podem ser substituídas por funções de apenas um argumento foi originalmente proposta por Gottlob Frege em 1879 (em seu primeiro grande texto de lógica, *Begriffsschrift*) e depois em 1920 pelo matemático russo Moses Schönfinkel (em uma palestra, de conteúdo somente publicado em 1924), um dos criadores da lógica combinatória. Atualmente o processo é denominado *currying* em inglês, em homenagem ao lógico norte-americano Haskell Curry, um continuador das ideias de Schönfinkel.

Este procedimento, em que operadores são definidos por meio do preenchimento de valências de outros operadores, ocupa uma posição central neste trabalho. Por exemplo, o operador verbalizador *V* é definido com várias posições argumentais em aberto; o verbalizador estativo *V_{BE}* e o verbalizador ativo *V_{DO}* são então definidos pelo preenchimento de valências de *V*. Assim, um operador é definido de modo geral e dá origem a versões especializadas por meio de determinadas parametrizações.

3.5 Assinaturas funcionais de operadores

Até o momento os valores com que lidamos são definidos como sequências de caracteres, ou *strings*; entretanto, as expressões linguísticas são usualmente mais complexas do que *strings*, como se observa de nomes, verbos, preposições etc..

Reduzir toda a gama de categorias gramaticais apenas a *strings* não apresenta dificuldades à versão do cálculo-lambda utilizada nos exemplos anteriores, em que não há qualquer re-

ferência explícita aos tipos dos valores. Temos utilizado o chamado *cálculo-lambda não-tipado* (*untyped lambda calculus* ou *type-free lambda calculus*), em que os operadores são definidos sem que os tipos das variáveis sejam formalmente especificados. Em certas situações, entretanto, pode tornar-se importante especificar os tipos dos parâmetros de um operador como parte da definição do próprio operador. Para tanto é preciso recorrer ao *lambda-cálculo tipado* (*typed lambda calculus*). Por exemplo, suponhamos que o operador *suffix* seja interpretado como a concatenação de uma *string* à direita de uma *raiz*, produzindo como resultado um *nome*. Sob esta interpretação, algo como *suffix* 'a' *menin* (em que 'a' e *menin* são, respectivamente, dos tipos *string* e *raiz*) gera o resultado 'menina' (de tipo *nome*).

O operador *suffix* pode então ser definido por uma expressão do lambda-cálculo tipado, de um modo em que os tipos de cada variável e do resultado computado fazem parte da definição do operador, como é mostrado abaixo:⁷

$$\begin{aligned} \text{suffix} &: \text{String} \rightarrow \text{Root} \rightarrow \text{Name} \\ \text{suffix} &= \lambda xy. \|yx. \end{aligned}$$

Nesta expressão:

- (a) $\text{suffix} : \text{String} \rightarrow \text{Root} \rightarrow \text{Name}$ é a assinatura funcional da operação;
- (b) $\text{suffix} = \lambda xy. \|yx$ é o corpo da operação; a definição da computação e da estrutura de variáveis da operação.
- (c) $\text{String} \rightarrow \text{Root} \rightarrow \text{Name}$ é o tipo da operação, em que *String* e *Root* são os tipos das variáveis ligadas *x* e *y*, respectivamente, e *Name* é o tipo do resultado computado pela operação.

A escolha dos tipos do exemplo acima tem caráter meramente ilustrativo; somente se podem atribuir tipos às expressões-lambda de uma teoria de acordo com os conceitos da própria teoria. No nosso caso, os tipos dos operadores e valores serão construídos ao longo de todo o texto. A análise mais formal do conceito de núcleo (seção 5.11, p. 59) recorrerá à tipagem do operador *h*, o núcleo universal.

⁷Os tipos serão denotados por identificadores iniciados por letras maiúsculas e nomeados preferencialmente em inglês.

Capítulo 4

Ideias fundamentais

4.1 Um sistema computacional único

A abordagem parte do pressuposto chomskiano de que a Gramática Universal possui um sistema computacional único: “...*existe apenas um sistema computacional e um [único] léxico* ...”, Chomsky (1993a, p. 170). Como sistema computacional, a GU compõe-se de *elementos* e de *associações* entre os elementos. Os elementos são de três tipos:

- (a) *Unidades portadoras de significado*, denominadas *traços semânticos*, como *fem*, *st* e *pl*. Na presente proposta, determinados traços semânticos são tomados como pertencentes à GU, como: *st*, estatividade; *ac*, atividade; *in*, incidência; *ec*, causalidade externa; *ic*, causalidade interna. Outros traços semânticos são considerados como específicos de (famílias de) línguas, como: *fem*, feminino; *masc*, masculino; indicativos de *humano* e *não-humano*; indicativos de *coisas redondas* e *coisas compridas*, etc. Existem também traços semânticos sobre os quais há dúvidas se pertencem à GU, como: *pl*, plural; *sing*, singular; *an*, animado; *ct*, contável. Provavelmente esses traços semânticos pertencem à GU, mas não será feito qualquer esforço para decidir a questão. O único ponto importante é que *todos* os traços semânticos podem ser manipulados pela GU.
- (b) *Unidades descritoras de expressão fonológica*, denominadas *traços fonológicos* como *a* ou *η*. Todos os traços fonológicos são considerados pertencentes à GU, embora, evidentemente, cada língua particular somente utilize parte deles.
- (c) *Operadores gramaticais*, como *r*, *n* e *v*, respectivamente o formador de raízes, nomes e verbos; *comp*, operador que associa uma raiz a um nome; *asp*, operador indicador da situação de um processo, como *em desenvolvimento* ou *concluído*, etc.. Todos os operadores gramaticais são tomados como pertencentes à GU.

Neste trabalho, manter-nos-emos restritos a um pequeno número de traços semânticos universais, os já citados *st*, *ac*, *in*, *ec* e *ic*, supostamente necessários à formação dos constituintes. Os traços poderão ser precedidos de sinais: +, indicativo de obrigatoriedade; ou -, indicativo de proibição. Assim: +*ic* significa existência obrigatória de causalidade interna; –*ec* indica a

proibição de causalidade externa. Caso o traço apareça destituído de sinal, ele é opcional (não-obrigatório, nem proibido); por exemplo: *in* indica que uma raiz pode ou não incidir sobre um outro termo. Uma hipótese fundamental é a de que um *elemento vocabular* (o constituinte mais básico de uma derivação) é necessariamente ou *estativo*, *+st*, ou *processual*, *+ac*,¹ mas não ambos simultaneamente.

Determinados traços semânticos não-universais recebem expressão gramatical (fonológica ou estrutural) geral e sistemática em línguas particulares, como *fem* em português. Traços deste tipo são comuns nos sistemas pronominais da vasta maioria das famílias linguísticas, e estão presentes, também, na expressão de modalidade de posse, como *alienável* ou *inalienável*. Tais traços serão denominados *traços semânticos gramaticalizados*. Não se pode prever que traços não-universais um idioma irá gramaticalizar. Os traços semânticos não-universais geralmente são implementados, translinguisticamente, por meio de *classificadores*, marcas morfosintáticas (como a alternância *o/a*, do português, para expressar os traços *masculino/feminino*, ou os classificadores nominais das línguas bantu).

Dependendo do idioma, outros traços semânticos não-universais recebem um tratamento que varia entre o gramatical “relativamente” *ad-hoc* (não completamente geral) e o não-gramatical completamente específico. É o caso: (i) de violações das regras gramaticais do caso anterior (como na alternância *avô/avó*, em português); (ii) da derivação vocabular peculiar, como em ‘abelha’/‘apicultor’; (iii) da atribuição de classificações como *felino/canino*; (iv) de informações pragmáticas (como o grau de periculosidade de um animal). Tais traços são costumeiramente denominados *traços enciclopédicos*, uma designação de origem semiótica, Eco (1991b). Os traços enciclopédicos refletem diretamente a história de cada língua, da comunidade e de falantes específicos, escapando ao que normalmente se considera um tratamento gramatical geral e sistemático.

A presente pesquisa contempla apenas os traços semânticos gramaticalizados, deixando à parte os traços enciclopédicos. Entretanto, o modelo proposto pode ser estendido, de modo natural, para o tratamento de ambos os tipos de traços.

Os elementos são associados por meio de operadores denominados *núcleos*.² Os núcleos estabelecem os ambientes para a formação de constituintes, essencialmente criando o cenário para a introdução dos operadores gramaticais. Por exemplo, o núcleo *N* dá origem a constituintes nominais, ao introduzir o operador nominalizador *n* em um constituinte previamente construído, desde que satisfeitas certas condições sob controle do próprio núcleo. Do mesmo modo, o operador *tense* é introduzido pelo núcleo *Tense*, responsável pela modalização temporal de um constituinte. Todos os núcleos pertencem à GU, embora cada língua particular selecione uma parte deles.

¹O símbolo *ac* foi adotado para seguir uma certa tradição; mas a interpretação é a de *processo*, mais geral do que a de *atividade/activity*. A interpretação dos traços *ec* e *ic* é a de *causalidade*, mais geral do que a de *agentividade*, e distinta da de *causa*, já consagrada ao termo que ativamente provoca a ocorrência de um evento, como ‘Maria’, em ‘Maria fez a criança dormir’.

²A formalização do conceito de núcleo é um dos temas da pesquisa.

O conjunto dos núcleos de uma língua compõe um dos subsistemas computacionais da GU, a *sintaxe*. Há pelo menos mais dois outros subsistemas computacionais da GU, a *semântica* e a *fonologia*. A pesquisa tem por objetivo mostrar que toda a realização argumental se processa na *sintaxe*.

A menos que se consiga mostrar que a *semântica* e a *fonologia* também possam ser desenvolvidas sobre o conceito de núcleo, esses dois sistemas computacionais ficam fora da *sintaxe* e, assim, dos objetivos do trabalho. É bastante provável que os métodos da *sintaxe*, independentemente de perspectiva teórica, sejam fundamentalmente inadequados à *semântica*. Pode haver uma certa similaridade fundamental entre a *sintaxe* e a *fonologia*, porém, o assunto foi deixado à margem neste trabalho.

Entretanto, considerar que a *sintaxe*, a *semântica* e a *fonologia* constituem sistemas computacionais distintos não significa dizer que a *sintaxe* não “enxergue” os resultados dos dois outros sistemas. Um ponto fundamental da proposta é que a formação vocabular ocorre na *sintaxe*. Virtualmente toda a formação de raízes, nos idiomas semitas, leva em conta combinações de traços fonológicos, processo em que a *sintaxe* interage com a *fonologia*. E absolutamente toda a formação das estruturas sintáticas baseia-se em restrições semânticas, que explicam, por exemplo, as semelhanças e diferenças estruturais entre ‘O repolho azedou’, ‘Maria adoçou o café’ e ‘O macaquinho engavetou-se’. Será mostrado que a formação destas e de outras estruturas deriva de associações específicas dos traços semânticos *st*, *ac*, *in*, *ec* *ic*, postas em combinação com determinados operadores gramaticais. As operações sintáticas se desenvolvem sobre um número bastante reduzido de traços semânticos, mas a eles recorrem de modo fundamental.

Sob a perspectiva adotada na pesquisa, a inserção do material semântico e fonológico ocorre já nas etapas iniciais da derivação sintática, e determina as possibilidades de formação dos constituintes.

As operações sintáticas selecionam e verificam determinados traços semânticos e fonológicos, mas não alteram tais traços, nem manipulam o seu conteúdo interno.

A única hipótese que se faz a respeito da organização dos traços semânticos e fonológicos, para que eles se tornem visíveis para a *sintaxe*, é a de que eles se organizam em *sequências*. Ou seja, a *sintaxe* processa sequências de traços semânticos, como $\langle +st, fem, sing \rangle$, e fonológicos, como $\langle f, \emptyset, 'n, \varepsilon, t, l, k, s \rangle$. Mas o processamento se restringe a verificar a boa formação da sequência e a existência de determinados traços. Como se verá, nenhuma outra operação é realizada, em decorrência da estrutura do núcleo.

De um modo geral, os traços semânticos serão denotados por $\sigma_1, \sigma_2, \sigma_3$, etc., e uma sequência de traços semânticos por $\langle \sigma_1, \sigma_2, \dots \rangle$, ou, de modo abreviado, $\langle \sigma_i \rangle$. Analogamente, os traços fonológicos serão geralmente denotados por φ_i e a sequência deles por $\langle \varphi_1, \varphi_2, \dots \rangle$ ou $\langle \varphi_i \rangle$. Uma sequência $\langle \sigma_i \rangle$ de traços semânticos pertencentes (internos) a um constituinte será denominada “o componente semântico” do constituinte. Cometendo um pequeno abuso de linguagem, o componente semântico também será denominado “a *semântica*” do constituinte. Do mesmo modo, a sequência $\langle \varphi_i \rangle$, interna a um constituinte, será o *componente fonológico* ou

“a fonologia” do constituinte.

Como veremos nos próximos capítulos, um núcleo categorizado como nome possui o formato $\lambda f.(f(x\langle\sigma_i\rangle))\langle\phi_i\rangle$, ao ser expresso como constituinte. Neste tipo de constituinte, x é um operador composto, formado pela aplicação sucessiva de operações da sintaxe. A estrutura do constituinte mostra um fato de extrema importância para a teoria da sintaxe, e que foi tema de análise no capítulo anterior: a variável f , pelo fato de encontrar-se não-especificada, impede que o operador $x\langle\sigma_i\rangle$ interaja funcionalmente com $\langle\phi_i\rangle$. Ou seja, impede que a semântica de um núcleo interaja com a sua fonologia. Como também foi visto anteriormente, a operação $x\langle\sigma_i\rangle$ também permanece em suspenso, sem que a semântica seja processada pelo operador x .

Esta é uma propriedade geral dos núcleos, e que será aqui reenunciada, em novos termos: os operadores sintáticos verificam “sequências” de traços semânticos e fonológicos, mas não alteram os traços individuais e não forçam cada traço específico a “produzir resultados”, o que somente ocorrerá em outros subsistemas da GU.

Por exemplo: o traço semântico *fem* manifesta-se fonologicamente em português de um modo padrão, sufixando a sequência fonológica (unitária) $\langle a \rangle$ a uma raiz nominalizada;³ porém, em associação com determinadas raízes, pode manifestar-se de outros modos, como no caso do feminino pronunciado como ‘avó’. Entretanto, ao longo de todas as operações sintáticas, a fonologia do traço *fem* jamais se realizará. É o que resulta da estrutura do núcleo. A realização fonológica somente ocorrerá após todas as operações sintáticas terem-se concluído, em uma etapa “pós-sintática” de realização fonológica.

Todos os fenômenos (morfo-sintáticos) de expressão fonológica fundamentam-se nesta propriedade dos núcleos, de processamento postergado da semântica e da fonologia. Em especial, toda a teoria da morfologia verbal e nominal, incluída a atribuição de caso, dela depende.

E esta é uma importante diferença entre a minha proposta e a da Morfologia Distribuída. Para obter o processamento postergado do material fonológico, a MD postula o princípio da “*late insertion*”. No modelo que proponho, o próprio sistema computacional assegura que não ocorra o processamento indesejado. Não é necessária, assim, a existência de um algoritmo de inserção de material fonológico em uma estrutura previamente derivada.

Voltemos, agora à estrutura que alegadamente possui um constituinte categorizado como nome:

$$\lambda f.(f(x\langle\sigma_i\rangle))\langle\phi_i\rangle.$$

Antecipando alguns resultados de capítulos vindouros, um *nome* é tipicamente formado a partir de uma *raiz*, a qual, por sua vez, é formada a partir uma entidade mais abstrata, o *elemento vocabular*, termo básico de qualquer derivação sintática. Deste modo, o operador x resulta da composição dos operadores *n*, *r* e *ve*. Deixando de lado algumas importantes análises ainda a

³Relembrar que traços semânticos não-universais, como *fem*, são interpretados como classificadores, de acordo com a conceituação de alguns parágrafos atrás: traços como *fem* e *masc* não implicam necessariamente que eventuais seres por eles caracterizados sejam dotados de determinadas qualidades, apenas indicam que a língua os classifica de um certo modo. Assim, não seria necessário dizer que “*mesa não é fêmea*” e que “*banco não é macho*”, apenas o faço para evitar possíveis incompreensões.

serem feitas, um nome possui, então, a estrutura

$$\lambda f.(f(n \circ r \circ ve \langle \sigma_i \rangle)) \langle \varphi_i \rangle,$$

em que x foi substituído pelo operador composto $n \circ r \circ ve$, e o símbolo \circ (também um operador) denota a composição de operadores.

Podemos, agora, apreciar o significado de um constituinte para os (pelo menos) três subsistemas computacionais da GU.

O operador $n \circ r \circ ve$ é um repositório de informações sintáticas, armazenadas individualmente em cada operador (n , r e ve); as informações sintáticas também possuem um efeito conjunto, decorrente da composição. Diremos que este é um *operador sintático*. O “resultado” do operador composto $n \circ r \circ ve$ é dotar o constituinte do comportamento sintático de um nome; ou seja, $n \circ r \circ ve$ é um categorizador (formador) de nomes. Deste modo, um nome é uma entidade complexa, formada na sintaxe, fato ainda a ser aprofundado. Também verbos são categorizados por operadores semelhantes a este; toda a sintaxe se processa pela ação de operadores que codificam informações sintáticas. O operador de um constituinte é o seu *componente sintático*; este é o único componente de um constituinte que expressa a sua lógica interna ao ser manipulado pela sintaxe.

O componente semântico, $\langle \sigma_i \rangle$, é um repositório de informações semânticas, parcialmente disponíveis para a sintaxe e para a fonologia, e que se constitui na principal entrada para o subsistema semântico (as outras duas entradas são o componente sintático e o fonológico). Segundo a presente proposta, o componente semântico obrigatoriamente contém traços semânticos universais e (não-universais) gramaticalizados, e, opcionalmente, uma referência a traços enciclopédicos; todos os traços são selecionados de acordo com a língua específica.

Analogamente, a sequência fonológica $\langle \varphi_i \rangle$ armazena informações fonológicas a serem operadas pelo subsistema fonológico, e apenas parcialmente pelos outros dois subsistemas de GU.

Portanto, um constituinte é um repositório de informações disponíveis para os subsistemas computacionais da GU. Caso seja válida a estrutura proposta para o núcleo, toda a formação do constituinte se processa necessariamente na sintaxe. Isto é: a sintaxe é o subsistema computacional da GU que prepara a informação a ser processada pelos demais subsistemas da GU.

Este resultado é compatível com a estrutura da gramática proposta por Chomsky (1981) e retomada por Chomsky (1995), em que a sintaxe gera estruturas a serem representadas em forma fonética (*phonetic form*) e em forma lógica (*logical form*), por módulos específicos da GU. Satisfaz-se assim o preceito de Chomsky (1995, p. 17): “No nível mais geral de descrição, o objetivo da gramática é expressar a associação entre representações de forma e representações de significado.”

Um constituinte $\lambda f.(f(w \langle \sigma_i \rangle)) \langle \varphi_i \rangle$ codifica uma associação particular, categorizada

por w , entre representações elementares de forma, $\langle \phi_i \rangle$, e de significado, $\langle \sigma_i \rangle$, situando w sob a ação de um contexto f não especificado. Potencialmente, o constituinte pode ser expresso como forma ou como significado, dependendo do operador atribuído a f (em uma etapa *pós-sintática*). Ou seja, o constituinte é a própria associação, gerada na sintaxe, entre o conteúdo de forma e o de significado de uma expressão linguística.

Um constituinte é uma estrutura computacionalmente em aberto, à espera de uma operação que preencha a variável f . Como veremos nos próximos capítulos, as operações sintáticas produzem constituintes como resultado, o que significa dizer que somente operações pós-sintáticas, de realização semântica ou fonológica, podem gerar um termo finalizado a partir de um constituinte.

Este trabalho tem como foco uma implementação computacional do subsistema sintático da GU, aplicada à formação das estruturas verbais e nominais, nos limites da realização argumental. Apenas marginalmente tocaremos os subsistemas semântico e fonológico, naquilo que diretamente contribuem para a formação das estruturas sintáticas.

4.2 Teorias da formação argumental

Um dos principais desafios da teoria linguística é explicar as diferentes estruturas argumentais a que um mesmo verbo pode dar origem, fenômeno conhecido como *realização argumental múltipla*.

As múltiplas realizações argumentais são caracterizadas por coleções de sentenças em que o mesmo verbo se encontra envolvido em diferentes configurações de argumentos. Uma das mais estudadas variações de estrutura argumental é a alternância *causativa/incoativa*:

- (1) Maria abriu a porta
- (2) A porta abriu

Esses exemplos ilustram a situação típica em que um *mesmo* conjunto de argumentos se alterna nas várias posições sintáticas disponíveis. Pertence a esta classe a alternância “*with/against*”, Levin & Rappaport Hovav (2005).

- (3) Kerry hit the stick against the fence
- (4) Kerry hit the fence with the stick

E também a alternância *dativa*, Harley (2002b):

- (5) John gave Mary a child
- (6) John gave a child to Mary

As duas últimas sentenças mostram que nem sempre a alternância dos *mesmos* argumentos preserva o significado das sentenças. Enquanto as sentenças com *hit* são paráfrases, o mesmo não se pode afirmar dos casos com *give*. Em 'John gave Mary a child' está implícito que Mary tem/terá um filho de John; já em 'John gave a child to Mary', nada se pode inferir além de que John entregou uma criança a Mary (Harley 2002b).

As estruturas em que os *mesmos* argumentos se alternam nas diferentes posições sintáticas, como as anteriores, têm sido objeto de pesquisa desde os primeiros anos da gramática gerativa e encontram-se razoavelmente bem resolvidas segundo diferentes linhas de pesquisa. Porém, os testes mais severos, sujeitos ainda a intensas pesquisas, são colocados pelos verbos que possuem múltiplas estruturas argumentais, como em Levin & Rappaport Hovav (2005):

- (7) Kim whistled
- (8) Kim whistled at the dog
- (9) Kim whistled a tune
- (10) Kim whistled a warning
- (11) Kim whistled me a warning
- (12) Kim whistled her appreciation
- (13) Kim whistled to the dog to come
- (14) The bullet whistled through the air
- (15) The air whistled with bullets

4.2.1 Papeis semânticos

Desde o período inicial da gramática gerativa, considera-se que as múltiplas estruturas argumentais têm origem no conteúdo semânticos dos verbos, em colaboração com outras partes da gramática: a entrada lexical de um verbo especifica a sua estrutura argumental. De um modo mais geral, o gerativismo chomskiano preconiza o princípio de que a informação armazenada na entrada lexical de um item determina o seu comportamento sintático. Esta ideia já está presente nas *relações temáticas (thematic relations)* de Gruber (1965), nos *esquemas de casos (case frames)* de Fillmore (1968), e, posteriormente, nas *grades temáticas (thematic grids)* de Stowell (1981). No trabalho pioneiro de Fillmore analisam-se as alternâncias

- (16) The door opened
- (17) John opened the door
- (18) The wind opened the door
- (19) John opened the door with a chisel

Fillmore (1968) propõe, pela primeira vez, a ideia de *papel semântico* (*semantic role*). Segundo Fillmore, o conteúdo semântico de um verbo é representado por uma lista de itens nominais em que cada elemento desempenha um papel, como *agente* (*agent*), *instrumento* (*instrument*), *tema* (*theme*) e *meta* (*goal*). Assim, o verbo *chutar* é caracterizado pelos papéis de *agente* e *tema*, e *enviar* por *agente*, *tema* e *meta*.

A teoria dos papéis semânticos possibilita uma explicação razoavelmente adequada das várias estruturas argumentais de um verbo como *open* acima. As múltiplas possibilidades de realização do sujeito são explicadas por meio de uma hierarquia de papéis, em que *agentes* precedem *instrumentos*, os quais, por sua vez, precedem *pacientes/temas*. Deste modo, um instrumento pode tornar-se sujeito, desde que falte um agente, e um paciente/tema pode ocupar a posição de sujeito, desde que faltem um agente e um instrumento.

Assim, a hierarquia de papéis temáticos explica a similaridade entre o sujeito de *abrir* intransitivo ('A porta abriu') e o objeto do mesmo verbo transitivo ('João abriu a porta'), e mostra também como um instrumento pode tornar-se sujeito ('A chave abriu a porta').

A ligação entre a descrição semântica de um verbo e as suas diferentes realizações argumentais dão origem ao chamado *linking problem*.⁴ As regras de associação entre as hierarquias temáticas e as estruturas argumentais representam a primeira tentativa de solução deste problema.

A abordagem de Fillmore, de associar grades temáticas a estruturas argumentais por meio de hierarquias temáticas, foi ampliada por Jackendoff (1972) e posteriormente fundamentou a pesquisa de numerosos outros estudiosos de diferentes orientações teóricas, entre os quais Givón (1984) e Van Valin (1993).

Chomsky introduziu os papéis temáticos na *Teoria da Regência e Ligação* (Chomsky 1981) ao dotar os núcleos lexicais de uma lista de papéis temáticos (Chomsky 1981, pp. 35–36) e ao postular uma correspondência um-a-um entre os papéis temáticos e os argumentos do núcleo, por meio do *Critério Teta* (θ -criterion): “Cada argumento porta um e apenas um papel- θ , e cada papel- θ é atribuído a um e apenas um argumento”, Chomsky (1981, p. 36).⁵ Chomsky também reformulou o *Princípio de Projeção* (Projection Principle) (Chomsky 1981, p. 29) em termos da teoria temática (Chomsky 1981, p. 38), de modo a refletir o “senso intuitivo” de que todo nível de representação sintática (LF e estruturas-S e -D) “deve ser uma projeção da estrutura temática e das propriedades de subcategorização das entradas lexicais ...”, Chomsky (1981, p. 36).⁶

Mark Baker, em sua tese de doutoramento (Baker 1985), orientada por Chomsky, formula o princípio UTAH (*Uniformity of Theta Assignment Hypothesis*), em que propõe que a

⁴A expressão *linking problem*, no restante do texto, será traduzida por *problema da associação* e não pelo mais natural *problema da ligação*, uma vez que o termo *ligação* possui um significado mais especializado em *Teoria da Regência e Ligação*, o que poderia conduzir a alguma ambiguidade.

⁵“Each argument bears one and only one θ -role, and each θ -role is assigned to one and only one argument.”

⁶“The natural intuitive sense of this principle is that every syntactic representation (i.e., LF-representation and S- and D-structure) should be a projection of the thematic structure and the properties of subcategorization of lexical entries ...”

correspondência entre os papéis temáticos e as posições sintáticas constitui um princípio universal, e é determinada pelo significado particular de cada verbo.

4.2.2 Estrutura de eventos

Atualmente, sintaticistas semanticamente orientados questionam se a informação codificada na entrada lexical de um verbo, ao estilo das grades temáticas, é suficiente para determinar a variedade de estruturas argumentais de que um verbo pode participar.

Desenvolveu-se, então, no âmbito gerativista, o enfoque alternativo de substituir a grade temática, uma estrutura de certo modo simples, por uma representação mais complexa, que codifica a semântica do verbo; a estrutura argumental sintaticamente realizada é então derivada desta representação por meio de um algoritmo. Ou seja, a representação sintática de um verbo é derivada de uma estrutura mais básica, que codifica a sua semântica lexical.

Tomaremos como exemplo desta abordagem a proposta de Levin e Rappaport Hovav (L&R), pela ampla fundamentação e por reunir contribuições de duas importantes propostas anteriores, as de Dowty (1979) e Vendler (1967).

A abordagem consiste em considerar que verbos descrevem eventos e que a entrada lexical de um verbo deve estar associada a uma descrição de evento. A questão, então, se reduz a como representar eventos, de modo que a estrutura dos eventos possa dar origem, por meio de um algoritmo apropriado, às diferentes estruturas argumentais dos verbos.

Levin & Rappaport Hovav (2005, pp. 71–73) analisam as representações semânticas de verbos pertencentes a três classes (verbos *causativos de mudança de estado*, *denominais*, e “*de modo*”):

(20) *dry*: [[x ACT] CAUSE [y BECOME <DRY>]]

(21) *bottle*: [[x ACT] CAUSE [y BECOME IN <BOTTLE>]]

(22) *jog*: [x ACT<JOG>]

Segundo esta teoria, verbos, como *dry* e *jog*, significam eventos “do mundo”; um evento possui uma estrutura que é representada por um predicado, como [[x ACT] CAUSE [y BECOME <DRY>]] ou [x ACT<JOG>]; um predicado é composto de predicados primitivos, como ACT, CAUSE, BECOME,

O primeiro exemplo mostra um verbo de mudança de estado em seu uso causativo, com a estrutura de predicado decomposta em predicados elementares que denotam as noções de causa e de mudança. Verbos deste tipo possuem a estrutura semântica

[[x ACT] CAUSE [y BECOME <STATE>]].

Os verbos que compartilham esta estrutura de predicado diferem apenas quanto ao estado espe-

cífico, <STATE>. Tem-se, assim, a separação entre o significado idiossincrático de cada verbo e a estrutura semântica, comum a toda uma classe de verbos.

Recentemente, como relatam L&R, a informação idiossincrática da estrutura de evento é tomada como sendo a “raiz”, de acordo com Pesetsky (1995). Segundo as duas autoras, o tipo ontológico da raiz, “*sua mais importante propriedade*”, amplamente determina a compatibilidade do verbo com o tipo estrutural de evento. Deste modo, então, deve ser interpretado o <STATE> do verbo *dry*: <DRY> é a raiz do verbo.

A estrutura do verbo denominal, *bottle*, deve ser interpretada analogamente. Deve-se observar que, neste caso, a noção de mudança é representada por BECOME IN, existindo assim um BECOME estativo e outro locativo.

No caso do verbo *jog*, a raiz é integrada à estrutura de evento de um segundo modo, como o “*modificador de um predicado*”, e anotada como um sub-escrito, <JOG>, à estrutura do evento.

Entretanto, como observam Levin & Rappaport Hovav (2005, p. 73), a estrutura de predicados, nos termos acima, parece ser insuficiente, como teoria geral, para explicar a estrutura argumental dos verbos. As autoras comparam duas classes de verbos de mudança de estado ou local: (a) *brake/quebrar* e *slide/deslizar*, que exibem alternância de transitividade, e (b) *murder/assassinar* e *remove/remover*, que não a exibem. Como, segundo L&R, tais verbos são dotados da mesma estrutura de eventos, a de *dry*, segue-se que a possibilidade ou impossibilidade de alternância de transitividade só pode advir de propriedades das raízes. L&R formulam então a hipótese fundamental de que o comportamento gramatical dos predicados verbais é determinado pela estrutura de eventos em combinação com propriedades semânticas da raiz: “*This property may reflect a general requirement on the pairing of roots with event structures that minimal elements of meaning encoded in the roots be given structural expression in the events structure*”.⁷ Ou seja, a estrutura argumental dos verbos é determinada pela estrutura dos predicados e pela semântica da raiz.

A representação de verbos por meio de estruturas de eventos oferece recursos mais abrangentes do que os proporcionados por listas de papeis semânticos, como os usados nas grades temáticas. Entretanto, muitas questões continuam em aberto.

A representação dos predicados evidencia problemas computacionais não adequadamente resolvidos:

- (a) A semântica múltipla de BECOME é resolvida por um recurso *ad hoc*, o acréscimo do qualificativo IN. Este recurso específico é insuficiente, por exemplo, para solucionar o denominar *entubar*, de ‘O médico entubou o paciente’/‘O paciente está entubado’, em português: o *paciente* não está *no tubo*, mas *o tubo* está *no paciente*. Seria então necessário algo como BECOME WITH? Quantas partículas a mais são necessárias para especificar a semântica de BECOME?

⁷“*Esta propriedade pode refletir um requisito geral sobre a associação de raízes com estruturas de eventos que a elementos mínimos de significado codificados nas raízes seja dada expressão estrutural nas estruturas de eventos*”. Grato à Maria Cristina Figueiredo por observações sobre esta tradução.

(b) O duplo modo como a raiz se insere na estrutura do predicado, como argumento de BECOME ou como subscrito de ACT, constitui um recurso notacional não motivado por razões conceituais, mostrando-se também uma solução *ad hoc*.

Tais expedientes mostram pontos em que a teoria é incompleta: os acréscimos introduzidos para viabilizar a computação dos casos particulares evidenciam lacunas conceituais, por não se basearem em princípios previamente elaborados e independentemente motivados.

Além dos problemas citados acima, de implementação operacional da teoria, há também problemas de fundamentos, como a determinação de um conjunto mínimo de predicados primitivos, sobre os quais não há consenso.

Tais observações, relacionadas a pontos em que a proposta de Levin e Rappaport Hovav é incompleta ou recorre a elementos *ad hoc* não pretendem desenvolver uma crítica de detalhes, mas pôr em discussão a abordagem como um todo.

Nos moldes de L&R, para abordar os problemas de estrutura argumental, devem-se desenvolver pelo menos três teorias: é necessário construir uma teoria semântica, uma teoria sintática, e uma teoria de projeção da semântica sobre a sintaxe (*linking theory*). Entretanto, a teoria semântica é pelo menos tão complexa quanto a teoria sintática que ela procura explicar. Ou seja, o peso do problema desloca-se da teoria sintática para duas outras teorias, uma semântica e outra de interface sintático-semântica.

As questões de determinação dos primitivos semânticos, de categorização dos primitivos lexicais (“*raízes*”) e das associações desses elementos entre si correspondem aos problemas inicialmente suscitados pelo comportamento sintático dos itens lexicais, sobretudo dos verbos. Reaparecem na semântica os problemas previamente surgidos na sintaxe, sem que métodos semânticos os tenham resolvido de modo satisfatório: as questões levantadas por estruturas como ‘O médico engavetou o livro’ (*livro posto na gaveta*) e ‘O médico entubou o paciente’ (*tubo posto no paciente*) não podem ser resolvidas pelo método de se agregar mais um qualificativo a BECOME.

Assim, os questionamentos dirigidos à proposta de L&R não são meramente episódicos: a decomposição semântica dos verbos busca emular o comportamento sintático destes elementos, e a harmonia buscada para uma teoria sintática não pode ser conseguida por meio de estipulações a uma teoria semântica, uma vez que o problema inicial, então, não estaria verdadeiramente resolvido, apenas mudaria de lugar.

Evidentemente, os problemas a serem resolvidos para a formação da estrutura argumental têm origem semântica, mas dificilmente a solução resultará de um tratamento apenas semântico, após o qual a representação sintática será obtida por consequência, por meio de um algoritmo de projeção semântico-sintático. Uma teoria da gramática, ao menos é o que creio, não pode separar em módulos independentes a construção semântica e sintática de um constituinte.

4.2.3 Hale e Keyser

Diferentemente das duas propostas anteriores, Hale e Keyser (H&K) partem do princípio de que a estrutura argumental de um item lexical resulta de suas propriedades estruturais, projetadas diretamente na sintaxe, e não do seu significado. Assim, a formação argumental não é resultante de estruturas semânticas (como as grades temáticas ou as estruturas de eventos) projetadas sobre estruturas sintáticas por meio de algoritmos especiais, mas de princípios que restringem as propriedades estruturais: *“During the course of our investigations, we have become persuaded that the proper representation of predicate argument structure is itself a syntax. That is to say, as a matter of strictly lexical representation, each lexical head projects its category to a phrasal level and determines within that projection an unambiguous system of structural relations holding between the head, its categorial projections, and its arguments (specifier, if present, and complement)”*, Hale & Keyser (1993, p. 53).⁸

Esta proposta é posteriormente aprimorada em:

(a) *“The term ‘argument structure’ is used here to refer to the system of structural relations holding between heads (nuclei) and their arguments within the syntactic structures projected by nuclear items.”*, Hale & Keyser (2001, p. 11).⁹

(b) *“These assumptions delimit a certain project: that of ascertaining the extent to which the observed behavior of lexical items is due to structural relations, rather than to the interaction of structure and some other component.”* Hale & Keyser (2002, p. 1).¹⁰

Considerada a época da proposta inicial (1993 e anos imediatamente anteriores), H&K ofereciam uma visão lexicalista alternativa à da teoria dos Princípios e Parâmetros de Chomsky (1981), eliminando a teoria temática como módulo independente da gramática. Em consequência, a proposta de H&K implicitamente questiona o próprio conceito de estrutura profunda (*deep structure*, DS). Deve-se observar que Hale & Keyser (1993) surge simultaneamente ao texto inicial da proposta minimalista de Chomsky (1993b), sendo ambos os artigos originalmente publicados em Hale & Keyser (Edts.) (1993). Assim, tanto H&K quanto Chomsky lançam, ao mesmo tempo, as bases para o questionamento da estrutura profunda como nível de representação linguística.

Para Hale & Keyser (2002, p. 13), há duas relações fundamentais de estrutura argumental:

⁸“Durante o curso de nossas investigações, nos tornamos persuadidos de que a representação adequada da estrutura argumental de um predicado é ela própria uma sintaxe. Isto significa dizer, no que concerne a uma representação estritamente lexical, que todo núcleo lexical projeta sua categoria sobre um nível frasal [sintagmático] e determina internamente a essa projeção um sistema não-ambíguo de relações estruturais existentes entre o núcleo, suas projeções categoriais, e seus argumentos (especificador, se houver, e complemento).”

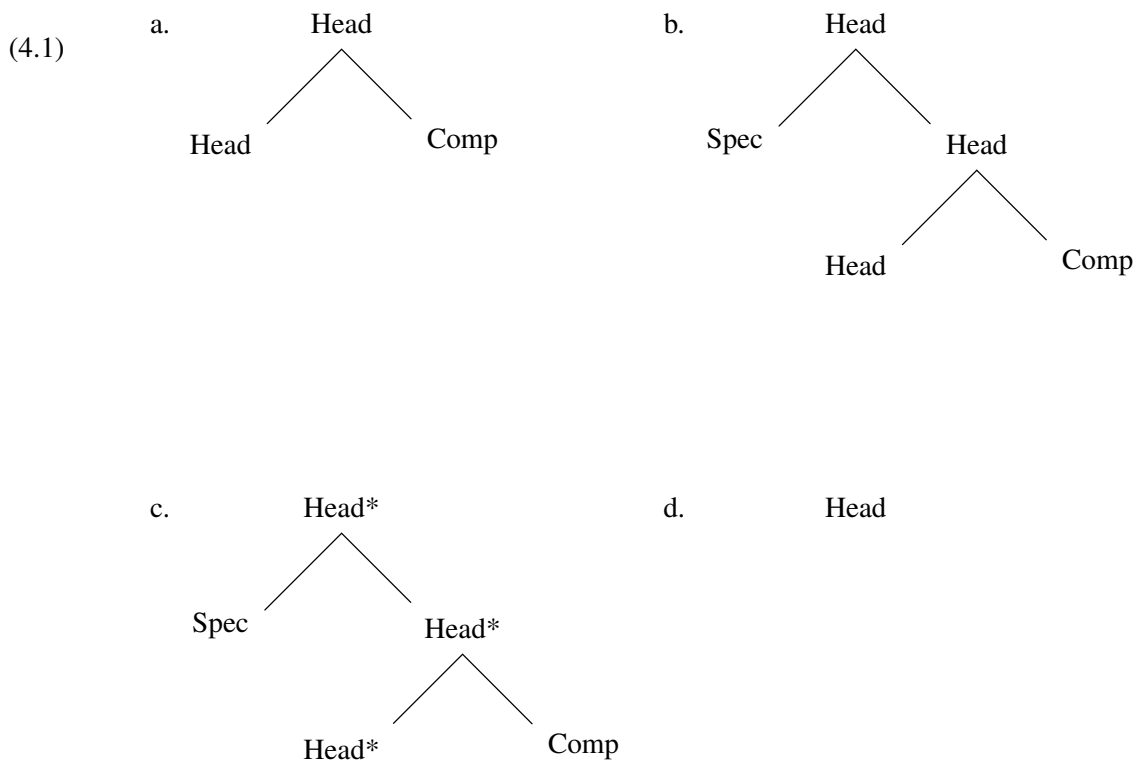
⁹“O termo ‘estrutura argumental’ é usado aqui para referir o sistema de relações estruturais existentes entre *heads* (nuclei) e os seus argumentos nas estruturas sintáticas projetadas por itens nucleares.”

¹⁰“Estas suposições delimitam um certo projeto: aquele de averiguar a extensão segundo a qual o comportamento observado dos itens lexicais se deve a relações estruturais, ao invés de à interação entre estrutura e algum outro componente.”

- (a) Núcleo-complemento (*head-complement*): Se X é o *complemento* de um núcleo H , então X é a única irmã de H (X e H c-comandam um ao outro);
- (b) Especificador-complemento (*specifier-head*): Se X é o *especificador* de um núcleo H , e se P_1 é a primeira projeção de H (i.e., H' , necessariamente não-vazia), então X é a única irmã de P_1 .

Assim, H&K desenvolvem a abordagem gerativista em que as relações sintáticas são estabelecidas por meio das relações entre núcleo, complemento e especificador.

As relações argumentais fundamentais, segundo H&K, permitem diretamente derivar as estruturas lexicais exibidas no próximo diagrama, em que “*Head* representa o núcleo e suas projeções categoriais, e *Comp* representa o complemento [e *Spec* representa o especificador]”. Tais estruturas lexicais serão por nós designadas *configurações estruturais básicas*:



As configurações estruturais básicas apresentam as seguintes características:¹¹

- (a) O núcleo toma um complemento mas nenhum especificador, projetando a estrutura denominada “monádica”;
- (b) O núcleo projeta uma estrutura composta de duas relações, núcleo-complemento e especificador-núcleo, formando o “tipo diádico básico”;

¹¹Nas análises do restante da seção, as letras a, b, c, d serão utilizadas para designar as configurações estruturais mostradas no diagrama.

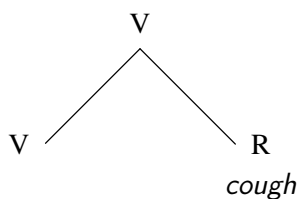
(c) O núcleo desta configuração “requer um especificador mas exclui um complemento”. O núcleo em questão é denotado por Comp, que projeta o especificador Spec com a ajuda de um outro núcleo, Head; a condição auxiliar de Head para a projeção de Spec é indicada pelo asterisco, tendo-se assim Head*;

(d) Um núcleo que não projeta complemento nem especificador, correspondendo ao “tipo atômico, o mais simples”.

De acordo com H&K, as configurações estruturais básicas são neutras com relação a categorias morfossintáticas (V, N, etc.): “*we will assume that morphosyntactic category and structural type are independent variables in the grammar of lexical projections*”, p. 14.¹² Entretanto, em inglês, predominantemente, mas não necessariamente, as estruturas realizam-se do seguinte modo (p. 13): (a) V; (b) P; (c) A; (d) N. Portanto: V projeta apenas complemento; P projeta especificador e complemento; A projeta especificador, mas somente com a participação de um outro núcleo; N não projeta estrutura argumental.

A forma como uma estrutura é pronunciada depende, em certos casos, da operação de *conflation*, “um processo importante da morfologia do inglês”: “... *the process according to which the phonological matrix of the head of a complement C is introduced into the empty phonological matrix of the head that selects (and is accordingly sister to) C*”, p. 12.¹³ Por exemplo, a fonologia do verbo *cough/tossir*, cuja derivação é parcialmente representada no diagrama a seguir, é obtida por *conflation* da matriz fonológica nula do núcleo V com a matriz fonológica da raiz *cough*. A operação de *conflation* é concomitante ao *Merge* (como esta operação é definida em Chomsky (1995)) do núcleo com o complemento.

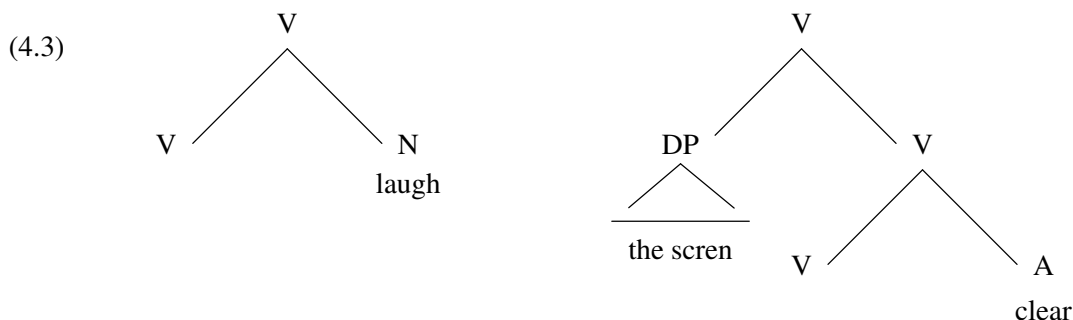
(4.2)



A proposta de H&K apresenta dificuldades para diferenciar a estrutura transitiva (lexical) da inacusativa. Para os dois autores, um verbo é um operador monádico (de tipo (a)), fato que permite a formação de verbos inergativos, como *laugh*, e incoativos alternantes, como *clear*, como é ilustrado no diagrama abaixo (Hale & Keyser 2002, pp. 15 e 16).

¹²“Assumiremos que categoria morfossintática e tipo estrutural são variáveis independentes na gramática das projeções lexicais”.

¹³“... o processo de acordo com o qual a matriz fonológica [componente fonológico] do núcleo de um complemento de C é introduzida na matriz fonológica vazia do núcleo que seleciona (e é portanto irmã de) C”.



A estrutura-base para derivação dos verbos incoativos-alternantes é de tipo (c), em que a posição de especificador é projetada pelo adjetivo, A, e não pelo núcleo verbal, V. Como V é de tipo (a), e a estrutura de tipo (b), em que o núcleo projeta Spec e Comp, é realizada por preposições,¹⁴ o paradigma de H&K não apresenta uma configuração estrutural básica adequada à formação das estruturas transitivas, fato que leva a um tratamento de certo modo complexo de tais estruturas. A este tema voltaremos posteriormente.

Fundamentalmente minha proposta difere da de H&K em dois importantes elementos: (a) a operação de *conflation* não ocorre de modo imediato, no exato momento do *Merge* do núcleo com o complemento; a formação fonológica, no modelo que proponho, é executada por um processo postergado, executado por operações pós-sintáticas; (b) *conflation* atua não apenas no caso de núcleos de fonologia nula; a operação se aplica de modo mais geral, com núcleos dotados de fonologia *defectiva*, isto é, que requerem complementação fonológica. Ademais, no modelo que proponho, um item lexical não projeta *diretamente* uma estrutura argumental, como as quatro configurações estruturais básicas de H&K; a formação argumental se passa de um modo mais indireto, como veremos no restante do trabalho — ficam assim solucionadas as dificuldades enfrentadas pelos dois autores para a representação das estruturas transitivas.

H&K são responsáveis por grandes avanços técnicos: propuseram uma implementação puramente sintática da projeção argumental de um item lexical. Ademais, mostraram, por meio da operação de *conflation*, que a formação da fonologia lexical de um item complexo (como os verbos denominais e tantos outros) se processa a partir da relação entre um núcleo e o seu complemento, seguida, exclusivamente, de operações fonológicas realizadas em uma cadeia de núcleos (por movimento ou, alternativamente, por cópia, uma questão em aberto).

A presente investigação se inscreve no programa de pesquisa inaugurado por Hale e Keyser; os resultados aqui apresentados muito devem às propostas dos dois autores.

¹⁴Embora H&K deixem em aberto a possibilidade de estruturas com Spec e Comp serem projetadas por quaisquer categorias lexicais, os dois autores efetivamente apenas consideram tais projeções como decorrentes de P, sobretudo para os dados do inglês. No estudo que fazem das estruturas transitivas (Hale & Keyser 2002, pp. 43–45), os verbos pertencem sempre ao tipo (a) e as preposições ao tipo (b).

4.2.4 Morfologia Distribuída

A Morfologia Distribuída (MD) alinha-se com o programa de Hale & Keyser (1993), Hale & Keyser (1998), não reconhecendo a especificação de papéis temáticos como propriedades dos itens lexicais. Assim, é tema de pesquisa na MD (Harley & Noyer April 1999) a definição configuracional dos papéis semânticos dos argumentos, em substituição aos papéis temáticos, compreendidos como elementos primitivos da teoria, como em Chomsky (1981). Neste sentido, a MD aborda a projeção argumental de um modo mais radical do que o fazem H&K, para os quais há dois sub-sistemas computacionais, um de derivação lexical e outro de derivação sentencial.

Na MD, não há a separação entre processos de formação lexical, morfológica e sintática, uma vez que as operações são efetuadas por núcleos pertencentes a um sistema computacional único. Não existem, assim, operações que ocorrem em uma fase pré-sintática, lexical.

As operações sintáticas da MD manipulam feixes de traços semânticos relevantes à formação lexical e sentencial, e não itens lexicais que contêm informação sintática, semântica e fonológica. O material fonológico é inserido nas estruturas sintaticamente derivadas em uma etapa pós-sintática, por meio de um algoritmo de inserção, segundo o processo de *inserção tardia* (*late insertion*).

Diferentemente da MD, minha proposta é lexicalista no sentido de Chomsky (1965b), p. 124–25: “*The lexicon is a set of lexical entries; each lexical entry, in turn, can be regarded as a set of features of various sorts. Among these are the phonological features and the semantic features ...*”.¹⁵ Por este modelo, os itens lexicais não definem grades temáticas, mas especificam traços fonológicos que são introduzidos na derivação juntamente com os itens lexicais. Neste sentido, o modelo que proponho é mais semelhante ao de H&K, com diferenças quanto aos mecanismos de inserção sintática dos constituintes.

Minha proposta também difere da da MD por não ser necessária a inserção tardia da fonologia; a estrutura de núcleo que proponho assegura o processamento postergado da fonologia gramatical e lexical.

¹⁵“O léxico é um conjunto de entradas lexicais; cada entrada lexical, por sua vez, pode ser considerada como um conjunto de traços de tipos variados. Entre esses [tipos] estão os traços fonológicos e os traços semânticos”

Parte III

Noção formal de núcleo

Capítulo 5

Formalização da noção de núcleo

5.1 A noção de núcleo

Toda a proposta sintática deste trabalho encontra-se fundamentada na noção de núcleo, único elemento que compõe uma derivação sintática. Analisaremos a seguir alguns importantes fenômenos captados pela noção, fenômenos a serem formalizados na próxima seção.

Consideremos palavras, da língua portuguesa, como 'estrutura', 'levedura', 'cobertura', 'lavratura' e 'brancura'. Tais palavras, segundo forte evidência, são formadas pela reunião de duas partes, como *estrut* e *ura*, para compor 'estrutura'. Uma parte, neste caso, aparentemente está ligada a algo estático, uma certa qualidade ou, até mesmo, uma coisa; outra parte, é uma sequência de sons, *ura*. Contudo pode-se dar por assegurado que o resultado da junção das duas partes é sempre um nome feminino: algumas vezes é um nome de coisa, como 'levedura'; outras, um nome de qualidade, como 'brancura'; em alguns casos, um nome de processo, como 'lavratura'; em outros pode pairar uma certa dúvida, como em 'cobertura'. Mas o resultado é sempre um nome, formado por dois termos postos em relação.

Diremos que o *processo gramatical* que reúne dois termos, cada termo de um determinado tipo, para formar um terceiro termo, de tipo também determinado, é um *núcleo*. Tem-se, assim, a abstração mais fundamental a ser captada pela noção de núcleo: *o núcleo é o processo responsável pela reunião de dois elementos*.

Não nos deteremos, neste momento, sobre o significado de *processo gramatical*, *termo* ou *elemento*; apenas buscamos caracterizar a ideia mais fundamental: *um núcleo estabelece uma relação binária entre termos previamente formados*.

Os termos postos em relação pelo núcleo denominam-se *argumentos* do núcleo. No caso recém-visto, diremos que o núcleo é um formador de nomes em *ura* e o formalizaremos por N_{URA} , utilizando uma notação que será intensivamente utilizada ao longo do texto. Segundo a visão desenvolvida neste trabalho, os nomes são formados por um certo núcleo, N , e este núcleo genérico pode ser especializado para determinadas finalidades, como é o caso de N_{URA} .

Como se pode observar, este núcleo não é visível no nome formado; nada há na pro-

núncia de ‘cobertura’ que mostre algo diferente das duas partes, ‘cobert’ e ‘ura’, exceto que as duas partes foram reunidas em uma certa ordem e formam um nome. O suposto núcleo N_{URA} só é detectado pelo resultado produzido; no caso, a formação de um nome de determinado tipo. Ou seja: *em geral, um núcleo não é diretamente percebido, somente se revelando pelos efeitos que causa*. Um núcleo é uma entidade *teórica, abstrata, não diretamente observável*.

Entretanto, o alegado núcleo N_{URA} é claramente inadequado para formar um nome por aplicação da sequência fonológica *ura* a um termo como *cas*, presente na formação da palavra ‘casa’. Cabe então ao núcleo, no caso, N_{URA} , estabelecer as condições para que dois termos possam ser aceitos como seus argumentos. Ou seja, não basta ao núcleo reunir dois termos, ele também precisa codificar regras de seleção adequada desses termos. Esta é outra função de um núcleo: *o núcleo contém uma regra de seleção de seus argumentos*.

Consideremos, agora, a palavra ‘meninas’, um *nome* que dá a ideia de seres humanos de pouca idade e do sexo feminino; poderia ser ‘meninos’, em que a característica masculina se revelaria pela vogal *o*. É, então, natural supor que existe uma certa base, *menin*, que “recebe”, nesta ordem, uma marca indicadora de feminino, *a*, e outra de plural, *s*, formando, assim, o nome pronunciado como ‘meninas’. O já mencionado núcleo N , ao formar o nome, é responsável pela aplicação das marcas de feminino e plural, ordenadamente, a um outro termo. As marcas podem ser indicadas pela sequência $\langle pl, fem \rangle$, o outro termo pode ser representado por *menin* (ao qual denominaremos *raiz*, um conceito ainda a ser formalizado), tendo-se então os argumentos do operador formador de nomes. O nome pronunciado como ‘meninas’ é, deste modo, o resultado da operação $N \langle pl, fem \rangle \text{menin}$, em que o operador N estabelece o cenário para que a sequência semântica $\langle pl, fem \rangle$ atue sobre a sequência fonológica *menin* e o resultado seja um nome que ao ser pronunciado o será como ‘meninas’. Tem-se aí uma outra propriedade do núcleo: *um núcleo é responsável pela interação entre seus dois argumentos*.

Neste trabalho, termos como *pl* e *fem* são denominados *traços semânticos* e são portadores de informação que contribui para o significado de outros termos. Existem traços semânticos que produzem consequências na fonologia, como *fem* e *pl*, e outros, como o traço estativo *st*, que não se revelam diretamente por marcas fonológicas (ao menos em português). O modo como a semântica se revela por meio da fonologia é próprio de cada língua, mas é sempre produzido por núcleos.

Há porém, um fato a considerar: a forma ‘meninas’, do português, encontra-se já “preparada” para ser pronunciada, após a aplicação dos traços semânticos à sequência fonológica. O mesmo, entretanto, não se pode dizer do exemplo análogo em latim: a interação entre o feixe de traços semânticos $\langle pl, fem \rangle$ e a base estativa de fonologia *puel* não dispõe ainda de toda a informação necessária para produzir a sequência fonológica a ser pronunciada; é necessária também a especificação de caso, que não se encontra disponível quando da aplicação dos traços semânticos à fonologia; ‘puellæ’ é apenas uma das possibilidades, sendo também possíveis outras formas, como ‘puellas’ e ‘puellarum’. No latim, como em muitos outros idiomas, não basta a interação entre a semântica e a fonologia da raiz para definir a forma finalizada de um nome

a ser pronunciado; são necessárias, também, informações sobre as relações existentes entre o nome e outros elementos da sentença a que o nome pertence. Portanto, em latim, se a semântica interagir com a fonologia, sem que esteja disponível a informação de caso, a sequência fonológica a ser pronunciada estará no mínimo incompleta, ou poderá ser considerada uma algaravia sem sentido. Uma vez que a semântica e a fonologia foram postas em contato por ação do núcleo *N*, cabe a este impedir que os dois termos por ele reunidos interajam “antes da hora” apropriada, ao menos em latim. Tem-se então outra característica dos núcleos: *um núcleo é responsável por postergar a interação entre os seus argumentos, até o momento em que estejam especificados os contextos em que a interação possa adequadamente ocorrer.*

O processamento postergado dos núcleos, muitas vezes, vai além da ação local que determina acréscimos fonológicos à forma como uma palavra é pronunciada. A organização das sentenças depende inteiramente desta propriedade. Em um par de sentenças como ‘O campo queimou’ e ‘A geada queimou o campo’, é visível a postergação. No momento em que o verbo (ou aquilo que se tornará o verbo) e o nome são postos em relação, não se encontra disponível ainda informação suficiente para definir o modo como o nome será interpretado, como sujeito ou como objeto. A relação entre o nome, *campo*, e o (futuro) verbo, *queimar*, pode ser interpretada de modo a que o nome atue como sujeito. É o caso da estrutura comumente chamada de inacusativa: ‘O campo queimou’. Mas, a mesma relação entre o nome e o (futuro) verbo pode ser expressa de um modo em que o nome é um objeto, como na estrutura transitiva ‘O geada queimou o campo’. O fato é que o contexto em que a relação será expressa (nos exemplos, inacusativo ou transitivo) não está disponível no momento em que o núcleo verbal e o nome são postos em contato. Não há sequer a pressuposição de que há um “núcleo verbal”: é possível também ‘queima do campo’. Mesmo no caso em que *campo* é interpretado como sujeito, a estrutura inacusativa é apenas uma das possibilidades; o termo pode eventualmente assumir a condição de sujeito de uma sentença passiva, ‘O campo foi queimado’, ou de sujeito de uma sentença média, ‘O campo queimou-se’. Em cada um dos casos, dependendo do idioma, pode haver consequências quanto ao modo como o nome e o verbo são pronunciados, decorrentes das relações estruturais entre o nome e o verbo. Toda a teoria gramatical depende desta propriedade de processamento postergado dos argumentos pelos núcleos. O conteúdo fonológico de um termo entra em uma derivação juntamente com o próprio termo, mas essa fonologia permanece “silenciosa” até ser situada em um determinado contexto e processada pelo núcleo adequado.

Assim, uma determinada combinação de núcleos estabelece o contexto para que uma associação específica de termos seja expressa como nome (‘Queima do campo’) ou como verbo (‘O campo queimou’, ‘O campo foi queimado’, etc.). Esta é uma importante propriedade dos núcleos: *os núcleos especificam contextos em que os seus argumentos são interpretados.*

Consideremos agora as sentenças ‘O campo queimou’, ‘O campo queimará’. Sabemos, como vimos acima, que a inacusatividade deriva de uma combinação específica de núcleos (um contexto; o modo como núcleos se combinam para dar origem a estruturas sintáticas é um dos temas deste trabalho). Como cada efeito gramatical é introduzido por um núcleo, haverá tam-

bém, nos exemplos dados, um núcleo que expressa o tempo, presente na forma dos verbos. Pelo contraste entre as sentenças, duas estruturas inacusativas que diferem apenas pela expressão do tempo, pode-se deduzir que o núcleo responsável por atribuir o tratamento temporal entra na derivação após a entrada dos núcleos que definem a estrutura como sendo do tipo inacusativo. Levando-se em conta o significado e a forma dos verbos, conclui-se que o núcleo atribuidor de semântica temporal se manifesta por meio de acréscimos morfológicos ao verbo ('queimou', 'queimará'). Esta é uma outra característica dos núcleos: *um núcleo é responsável por acréscimos semânticos e fonológicos a seus argumentos*.

Entretanto, o núcleo de tempo não se revela apenas na forma do verbo. Em 'O campo vai queimar', o tempo se exprime por meio de um auxiliar, 'vai'. Portanto, o núcleo atribuidor de tempo pode se manifestar morfológicamente, pela flexão do verbo, ou sintaticamente, pela introdução de um auxiliar em uma construção particular (no exemplo, *auxiliar no presente + infinitivo*). Assim, os acréscimos fonológicos de um núcleo podem manifestar-se morfológicamente, sintaticamente ou por uma combinação de expressões morfológicas e sintáticas. Portanto, *o mesmo núcleo pode se manifestar de diferentes modos, dependendo dos contexto em que se encontra inserido*.

O ponto fundamental é que *a expressão de um núcleo não se dá de modo isolado, mas sim em decorrência da combinação com outros núcleos*. Tal fenômeno não é exclusivo das estruturas verbais. Em ambiente nominal, um exemplo da expressão dependente de contexto é o da relação genitiva em inglês, expressa sob, pelo menos, dois formatos, por meio de uma preposição ('Book of Mary'), ou de uma desinência ('Mary's book'). A forma em que a relação genitiva é pronunciada em inglês, desinência ou preposição, depende de alguma nuance de significado ou de um efeito estilístico. De todo modo, trata-se de um acréscimo semântico codificado por um núcleo.

A expressão contextual de núcleos, um dos fatos mais fundamentais da gramática, pressupõe um outro fato igualmente fundamental: ao entrar em uma derivação, um núcleo precisa preservar a sua individualidade sintática; não pode ser "absorvido por" nem "diluído em" outro núcleo. Não há como antecipar os contextos de que um núcleo pode participar, *portanto um núcleo deve preservar a sua individualidade sintática ao longo de toda a derivação*.

Em decorrência da expressão contextual, um falante pode ter opções para expressar uma combinação de núcleos. O futuro de uma sentença ativa, por exemplo, pode ser expresso de vários modos: 'Maria vai falar', 'Maria falará', 'Maria irá falar' 'Maria fala amanhã'. Pode haver diferenças sutis de significado nas formas escolhidas, mas o fato é que o tempo verbal em PB pode ser expresso diretamente no verbo, por meio auxiliares, ou de advérbios. Em outros casos, a possibilidade de escolha é bem menor: o genitivo em português é virtualmente restrito a construções com a preposição *de*. É muito questionável se os vários modos de expressar o tempo verbal são todos diferentes expressões da mesma estrutura. De todo modo, com ou sem possibilidade de escolha, o falante, para expressar uma combinação de núcleos, sempre terá que selecionar uma opção oferecida pelo sistema da língua; não lhe é dado alterar o modo como a

língua forma plurais ou expressa os verbos, por exemplo. Tem-se assim outra propriedade dos núcleos: *um núcleo é portador de expressão própria*,¹ determinada pelo sistema da língua, fato habitualmente denominado *expressão gramatical* do núcleo.

Por fim, os diferentes modos de atuação de um núcleo, acima sumarizados, são regidos por regras especificadas no interior dos próprios núcleos, quando então tem-se que *um núcleo é portador de uma lógica própria que rege a sua atuação interna e sua interação com outros núcleos*.

A discussão precedente a respeito das propriedades dos núcleos passou-se em termos relativamente intuitivos, uma vez que o tratamento preciso do tema depende da própria formalização do conceito, o que será feito a seguir.

5.2 Paradigma de representação

A noção de núcleo desta pesquisa é extremamente abstrata, uma vez que tem por objetivo formalizar os fenômenos da estruturação argumental com base neste único conceito sintático.

Na seção anterior desenvolvemos a noção intuitiva de núcleo, para dar conteúdo à formalização; nesta seção passaremos a construir, passo a passo, o conceito formal, procurando mostrar como cada ideia computacional utilizada tem em vista o problema de como simular o fenômeno da linguagem. Caso a formalização fosse apresentada diretamente, sem preparação, haveria o risco de ser considerada arbitrária e destituída de conteúdo empírico.

Sabemos que um núcleo é responsável por estabelecer uma relação entre dois termos. Este núcleo geral, que independe de quais termos são relacionados, será denotado por h (*head*). Um exemplo claro de núcleo é a preposição, um termo que liga dois nomes. A preposição *de*, em um caso particular, pode ligar *casa* e *Maria*, para formar a expressão pronunciada como ‘Casa de Maria’. Para utilizar uma notação mais formal, a preposição *de* será representada por P_{DE} ; sua aplicação aos dois nomes será, então, $P_{DE} Maria casa$. Ocorre que nossos núcleos serão sempre expressos por meio do núcleo genérico h , e neste momento, por falta de outros elementos, somente podemos identificar diretamente P_{DE} com h ; ou seja, $P_{DE} = h$. Assim, a aplicação da preposição aos seus argumentos pode ser expressa de dois modos, como é indicado pela igualdade $P_{DE} Maria casa = h Maria casa$.²

Em uma seção anterior já vimos o núcleo formador de nomes em ‘URA’. Analogamente à solução da preposição, com nossa solução provisória, a formalização deste núcleo em termos de h é imediata: $N_{URA} = h$ e, assim, $N_{URA} levedura = h levedura$, que produz um nome a ser posteriormente pronunciado como ‘levedura’.

A comparação de P_{DE} com N_{URA} coloca de pronto uma questão: como diferenciar o “ h

¹ Ainda que contextualmente determinada.

² *Maria* é o argumento interno e *casa* é o argumento externo do núcleo; a ordem dos argumentos é determinada pelo modelo computacional, tema a ser desenvolvido no restante da seção.

da preposição” do “*h* do nominalizador”? De certo que há diferença, pois nominalizadores são muito diferentes de preposições. Ou seja, *h* precisa ter uma estrutura interna que lhe permita formalizar situações tão variadas. A estrutura interna de *h* é um dos objetos da pesquisa, e começaremos a tratar disso algumas linhas adiante.

Há, porém, um outro problema, de ordem mais imediata, a ser resolvido. O nominalizador N_{URA} *sempre* acrescenta a sequência fonológica *ura* ao seu segundo argumento, e a solução $N_{\text{URA}} = h$, até aqui adotada, não permite representar o fato geral de que o primeiro argumento de N_{URA} é, necessariamente, *ura*. A notação possibilita definir valores particulares de núcleos e argumentos, mas não permite lidar com leis gerais envolvendo esses elementos. Ou seja, o núcleo, *h*, carece de qualquer capacidade de abstração. Além disso, até o momento, apenas temos dito que *h* possui dois argumentos, mas não há qualquer mecanismo que garanta tal condição.

Existem, assim, alguns problemas a resolver: assegurar que *h* seja (a) um operador binário, (b) dotado de capacidade de abstração, e (c) possuidor de uma estrutura interna.

Nossa abordagem desses problemas recorre ao λ -cálculo,³ um formalismo especialmente adequado à definição de modelos processuais abstratos, utilizado em vertentes da ciência da computação e em algumas linhas de pesquisa da semântica formal e da linguística. O λ -cálculo é o paradigma de formalização desta pesquisa; diagramas também serão extensivamente utilizados, para facilitar a visualização e ilustrar os aspectos intuitivos das análises.

5.3 Formalização do núcleo universal

Assegurar que *h* seja um operador binário se consegue com a definição

$$h = \lambda i e. i e.$$

A definição diz que *h* é “a mesma coisa” que a expressão $\lambda i e. i e$: em qualquer expressão em que se use *h*, pode-se alternativamente usar $\lambda i e. i e$. A expressão- λ especifica a estrutura do operador *h*: *h* possui exatamente dois parâmetros, representados pelas variáveis *i* e *e*; *i* e *e* são dois “guardadores de lugares” para termos a serem inseridos na expressão. As variáveis poderiam ter outros nomes, como *x* e *y*; os nomes escolhidos têm a evidente conotação de *interno* e *externo*. A ordem dos parâmetros, na expressão lambda, é importante: a expressão $\lambda i e$ define que o operador *h* primeiramente se aplica a um valor que ocupará a(s) posição(ões) indicadas por *i*, e somente depois é que se aplicará a um valor que preencherá a(s) posição(ões) indicadas por *e*. Diremos, então, que *i* recebe o *argumento interno* do núcleo e *e*, o *argumento externo*. Por abuso de linguagem, algumas vezes se diz que *i* e *e* são os argumentos interno e externo do núcleo, embora sejam variáveis e os argumentos sejam valores. Os argumentos interno e externo de um núcleo correspondem, muito proximamente, aos conceitos de *complemento* e *especificador* da

³Ou “cálculo lambda”; a forma λ -cálculo será preferencialmente usado no trabalho.

gramática gerativa. Portanto, h é um operador binário dotado de capacidade de abstração, dada a sua definição como expressão- λ , uma vez que i e e são termos não especificados.

Até o momento, a definição de h apenas contém referências aos argumentos do núcleo. Somente se tem i “processando” e , que é o sentido da expressão ie ; i denota um operador que se aplica a e .

O núcleo, definido deste modo provisório, já permite especificar uma parte da lógica de um dos nominalizadores citados na seção anterior, o que aplica os traços semânticos $\langle pl, fem \rangle$ à sequência fonológica *menin*, para produzir o resultado *meninas*. Em termos mais formais, a aplicação deste nominalizador aos seus argumentos pode ser expressa por $N \langle pl, fem \rangle menin$. Assim,⁴

$$N \langle pl, fem \rangle menin = h \langle pl, fem \rangle menin.$$

Como $h = \lambda ie. ie$, substituindo-se h pela sua expressão- λ equivalente, tem-se

$$N \langle pl, fem \rangle menin = (\lambda ie. ie) \langle pl, fem \rangle menin.$$

Na igualdade acima, o operador $\lambda ie. ie$ se aplica ao primeiro valor, o feixe de traços semânticos $\langle pl, fem \rangle$; este feixe, assim, substitui i , produzindo o resultado

$$N \langle pl, fem \rangle menin = (\lambda e. \langle pl, fem \rangle e) menin.$$

Tem-se, então, um novo operador, $\lambda e. \langle pl, fem \rangle e$, que se aplica à sequência fonológica *menin*. Substituindo-se a variável e pela sequência fonológica, resulta

$$N \langle pl, fem \rangle menin = \langle pl, fem \rangle menin,$$

Ou seja, a nominalização $N \langle pl, fem \rangle menin$ se reduz a uma aplicação, ordenada, dos traços semânticos $\langle pl, fem \rangle$ à fonologia *menin*. Em primeiro lugar, o traço semântico *fem* se aplica a *menin*; o resultado é a sequência fonológica *menina*; em seguida, *pl* se aplica ao resultado *menina*, acrescentando-lhe a marca do plural, obtendo-se, então, *meninas*. Portanto

$$N \langle pl, fem \rangle menin = meninas.$$

Vejamos aonde nossa definição do núcleo, $h = \lambda ie. ie$, nos levou. Como i , no exemplo acima, é um feixe de traços semânticos, $\langle pl, fem \rangle$, e e é um feixe de traços fonológicos, *menin*, a definição

⁴Na expressão $h \langle pl, fem \rangle menin$, os traços semânticos *pl* e *fem* são operadores que se aplicam a uma sequência de traços fonológicos, *menin*, produzindo outra sequência de traços fonológicos como resultado, *meninas*. Por este motivo, $\langle pl, fem \rangle$ é o argumento interno de h , enquanto *menin* é o argumento externo; relembrar que a ordem de processamento dos argumentos é *ie*, ou seja, i processa e (a semântica processa a fonologia, neste caso). A ordem de processamento é dada pela sequência $\langle pl, fem \rangle$, em que *fem*, por contiguidade, se aplica em primeiro lugar à sequência fonológica.

produziu um processamento em que a semântica opera sobre a fonologia:

$$\langle pl, fem \rangle \text{menin.}$$

Ora, em português, este processamento, em que a semântica produz efeitos fonológicos, pode ser considerado aceitável: o resultado *meninas*, uma sequência de traços fonológicos, será pronunciado 'meninas', seja como sujeito, objeto, ou argumento interno de preposição. Isto é: a fonologia produzida não será alterada pela atribuição de caso, em português. Mas, em latim, a situação análoga é diferente: a fonologia resultante não pode ser considerada completa, pois a morfologia da palavra se encontra em aberto, à espera de acréscimos fonológicos indicadores de caso, atribuídos pela estrutura verbal ou por uma preposição. Nas línguas da família tupi a situação é ainda mais problemática, uma vez que um nome em posição de objeto pode até mesmo ser absorvido pela raiz de um verbo, em um processo denominado *incorporação*.⁵

Ou seja, não se pode permitir que a semântica interaja prematuramente com a fonologia. Ao longo de uma derivação sintática, não é possível prever o tipo de ambiente em que um termo será incluído, e as consequências que daí advirão para a sua expressão fonológica. Além do mais, as soluções devem ser adequadas à gramática de qualquer língua humana. Portanto, os núcleos têm que manter a sua integridade sintática ao longo de toda a derivação. O nominalizador do nosso exemplo, *N*, não pode diluir-se em uma expressão fonológica, *meninas*, ao menos precocemente.

A necessidade de se impedir a interação entre os argumentos interno e externo de um núcleo não ocorre apenas nas situações de contato entre a semântica e a fonologia. É uma característica geral das línguas humanas, em que a relação entre dois termos é sempre interpretada segundo um contexto mais abrangente. Todos os fenômenos de formação das estruturas argumentais giram em volta deste princípio, que, em essência, significa: os termos são postos em relação de acordo com regras; mas, essas regras ainda não se encontram especificadas no momento em que a relação se estabelece.

Portanto, na definição de núcleo, $h = \lambda ie. ie$, é necessário interromper a interação entre as variáveis *i* e *e*.

Uma maneira de se impedir que os dois argumentos entrem em contato é recorrendo-se à noção de *pair* (Barendregt 1990, Hankin 2004):

$$pair = \lambda f.(f i) e.$$

No *pair*, a variável *i* é impedida de se aplicar à variável *e* pela variável *f*; e a variável *f* encontra-se em suspenso, pela ação de λf .

Dito de outro modo, mais adequado à nossa finalidade: as variáveis *i* e *e* encontram-se em relação, porém em uma relação não especificada, regida por *f*, um contexto ainda a ser

⁵Ver Baker (1985), o texto de referência para a noção de incorporação, e Hale & Keyser (2002) para o tratamento conjunto dos conceitos de incorporação e de *conflation*, este último de utilização futura em nossa proposta.

definido.

Tem-se então mais um passo em direção à definição do núcleo universal h , ao se substituir ie , a aplicação imediata do argumento interno ao externo, por $\lambda f.(fi)e$, que é a mesma aplicação, porém postergada por f :

$$h = \lambda ie.(\lambda f.(fi)e).$$

A definição acima ainda é incompleta, mas já captura uma importante propriedade das línguas humanas: o processamento contextual dos argumentos interno e externo de um núcleo (processamento de i sobre e contextualizado por f).

Abordemos agora a questão das exigências que um núcleo costuma impor a seus argumentos. Por exemplo, o argumento externo de um núcleo verbal deve ser nominal; uma raiz, como veremos, reúne um feixe de traços semânticos e um feixe de traços fonológicos. Todo núcleo espera aplicar-se a argumentos de um determinado tipo. É necessário então equipar a definição de h com um critério de seleção dos argumentos interno e externo. Faremos isto por meio de uma função booleana, b , que introduz uma lógica interna a h . O núcleo é agora definido por:

$$h = \lambda bie. \begin{array}{ll} \lambda f.(fi)e & \text{if } bie \\ error & \text{otherwise} \end{array}$$

Primeiramente, é preciso observar que a função b agora parametriza h , fato resultante da expressão λbie . A lógica interna de h funciona do seguinte modo:

- (a) a função b testa os valores atribuídos a i e e , para verificar se apresentam as características requeridas;
- (b) se o teste produzir o resultado *verdadeiro*, isto significa que os valores passados para i e e foram aceitos e o núcleo pode ser constituído;
- (c) neste caso, o núcleo assume o valor $\lambda f.(fi)e$, o resultado esperado;
- (d) se o teste produzir o resultado *falso*, o que caracteriza a condição *otherwise* acima, ocorre uma condição de erro; o núcleo recusou pelo menos um dos valores atribuídos a i ou a e ; e a montagem do núcleo falha.

Para exemplificar, suponhamos que b especifique que i deve ser preenchido por um feixe de traços semânticos e e por uma sequência fonológica, então a aplicação $b \langle pl, fem \rangle \text{menin}$ dará o resultado *verdadeiro*; já a aplicação $b \text{demenin}$ produzirá o resultado *falso*.

A definição de núcleo, do modo como se encontra especificada, é *genérica*, apresenta *processamento postergado*, realiza a *seleção de argumentos*, mas ainda não permite caracterizar uma preposição ou um verbo.

Partimos do pressuposto de que nomes, verbos, preposições, raízes etc., são construídos por operadores específicos, sempre binários. Assim, as variáveis i e e já compõem o cenário para que esses operadores sejam acrescentados à definição. Acrescentaremos então o operador denotado pela variável k , para realizar a operação específica, característica, do núcleo; k é introduzido de modo a tomar i e e como argumentos, e a preservar a propriedade de interação postergada entre estes argumentos. Deste modo, h é definido por:

$$h = \lambda bkie. \begin{array}{ll} \lambda f.(fki)e & \text{if } bie \\ error & \text{otherwise} \end{array}$$

O operador k realiza a *lógica gramatical* própria do núcleo, responsável pelo processamento dos argumentos i e e . E, além desta atividade interna, k , como qualquer termo, é também portador de informação codificada em sua estrutura, a ser processada por outros núcleos que entrarem acima dele em uma derivação.

Agora é possível especificar o que se pode chamar de “categoria gramatical” das preposições. Para fazê-lo, na estrutura de parâmetros $\lambda bkie$, devem-se suprir valores adequados para a função de seleção, b , e para o operador característico, k . O operador característico das preposições será denotado por p , que contém a lógica particular de uma dada língua para o tratamento geral de suas preposições. A função de seleção, em uma primeira aproximação, imporá a condição de que os argumentos do núcleo sejam nominais. Indiquemos tal função por β . Se designarmos a categoria das preposições por P , a especificação de P será então:

$$P = h\beta p.$$

Ou seja, h , que possui a estrutura de parâmetros $\lambda bkie$, aplica-se aos valores β e p . Assim, as duas primeiras posições argumentais de h (b e k) ficam preenchidas, e h torna-se um operador em que apenas as posições argumentais i e e permanecem em aberto.

Portanto,

$$P = h\beta p = \lambda ie. \begin{array}{ll} \lambda f.(fpi)e & \text{if } \beta ie \\ error & \text{otherwise} \end{array}$$

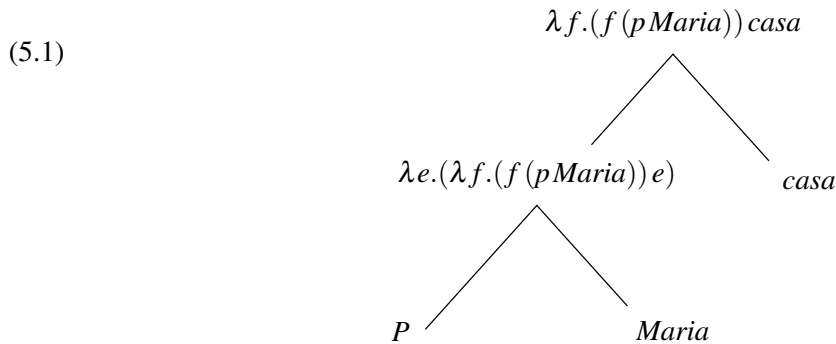
O operador gramatical P é o núcleo h parametrizado por β e p , possuindo a estrutura de parâmetros λie ; ou seja, é um operador cujas valências em aberto se reduzem aos argumentos interno e externo. Portanto, P , um operador binário, está preparado para formar uma estrutura preposicionada, desde que aplicado aos valores adequados.

Podemos agora representar a expressão ‘Casa de Maria’, em que ‘Casa’ é o argumento

externo da preposição e 'Maria' é o argumento interno, por meio de P :⁶

$$P Maria casa = (\lambda ie. (\lambda f. (f p i) e) \dots) Maria casa$$

Na igualdade acima, por motivo de clareza, a lógica de seleção foi deixada subentendida, representada por \dots . A aplicação de P aos dois valores pode ser representada por um diagrama:



O diagrama exhibe o que já constitui uma primeira derivação sintática, ainda que simples: P aplicado a dois valores nominais. Na raiz da árvore de derivação está a estrutura preposicionada obtida como resultado:⁷

$$\lambda f. (f (p Maria)) casa.$$

Este primeiro exemplo proporciona a oportunidade de algumas convenientes definições:

- (a) O operador gramatical P é um “núcleo parametrizado”, uma vez que é definido por uma parametrização especial de h , $h \beta p$. Para simplificar, entretanto, P será chamado apenas de “núcleo P ”. Este abuso de linguagem, tendo em vista que h é o único núcleo do sistema computacional, será também empregado com relação aos demais operadores gramaticais: núcleo N , núcleo V , etc..
- (b) A aplicação de um núcleo (parametrizado) aos argumentos interno e externo será denominada *derivação*. Como um argumento também pode ter sido obtido por derivação, o conceito é recursivo: uma sucessão de derivações também é uma derivação.⁸
- (c) No topo do diagrama anterior situa-se uma estrutura que é resultado de uma derivação: o resultado de uma derivação será chamado de *constituente*; se dirá que o constituinte foi *gerado* pela derivação. No exemplo dado, *Maria* e *casa* também são constituintes, por terem sido gerados em outras derivações, não exibidas no diagrama. Entretanto, P não é um constituinte, por não ter sido gerado em uma derivação; também o operador p , incluído pelo núcleo, não é

⁶Em termos gerativos chomskianos, ‘Maria’ é o complemento da preposição e ‘Casa’ é o especificador.

⁷Como a raiz da árvore de derivação se situa na parte de cima do diagrama, por comodidade, será usada, no restante do trabalho a expressão “topo do diagrama”.

⁸A possibilidade de uma derivação D aplicar-se ao resultado de outra derivação é controlada pelo critério de seleção do núcleo que gerou D .

um constituinte.

(d) O constituinte situado no topo da derivação, $\lambda f.(f(pMaria))\text{ casa}$, possui como único argumento a variável f . Isto significa que as valências correspondentes aos argumentos interno e externo foram previamente preenchidas. Um constituinte que possua apenas a variável f como valência disponível será denominado *constituinte maximal*. Um constituinte maximal será, muitas vezes, abreviadamente designado pela forma xP , em que x é o operador característico do núcleo; no exemplo em análise tem-se um pP .⁹ Já o constituinte $\lambda e.(\lambda f.(f(pMaria))e)$ não é maximal, por apresentar a valência e disponível.

O operador característico, k , já dota o núcleo, potencialmente, de toda a funcionalidade necessária à definição dos operadores gramaticais. Entretanto, sob a forma pouco desenvolvida em que se encontra, ele contém veladamente alguns elementos que devem ser explicitados sob a forma de parâmetros do núcleo.

Para aprofundar a análise do operador k , voltemos ao núcleo formador de nomes em *ura*, um nominalizador. Os nominalizadores em geral podem aplicar-se a raízes. Como uma raiz compõe-se de um feixe de traços semânticos e de um feixe de traços fonológicos, a parametrização do núcleo nominalizador, a que denotaremos por N , é imediata: ele possui uma função de seleção, β , que seleciona, como parâmetros interno e externo, os feixes de traços pertencentes a uma raiz, e é dotado de um operador característico, que designaremos por n . Assim:

$$N = h\beta n.$$

Porém, o nominalizador genérico acima, para o nosso propósito, precisa ser especializado para incluir a fonologia *ura* aos termos a que se aplica. Na realidade, a inclusão não é apenas da fonologia; devemos lembrar que os nomes a serem formados são todos classificados como femininos. Portanto, cabe a N fazer também uma inclusão de semântica (gramatical).¹⁰ Ou seja, o nominalizador em *ura* realiza acréscimos semânticos e fonológicos aos termos aos quais se aplica, além de nominalizar. Evidentemente, o acréscimo semântico é $\langle fem \rangle$ e o fonológico é *ur*.

O acréscimo de elementos semânticos e fonológicos, por parte de nominalizadores, é um fenômeno de ampla difusão translinguística. Aparece, por exemplo, também na formação de nomes abstratos, em *eza*, e de nomes de ação, em *ção*; que em português se realizam como sufixos. O fenômeno não é exclusividade da nominalização; a formação de verbos causativos em *izar* mostra o processo atuante no campo verbal. Advogamos, em capítulo vindouro, que as preposições particulares se caracterizam por acréscimos semânticos e fonológicos ao núcleo P .

⁹Um constituinte maximal corresponde a uma projeção máxima da Gramática Gerativa. Neste trabalho, não se pode usar a expressão “projeção máxima”, porque, como veremos, um núcleo não projeta argumentos.

¹⁰Segundo a proposta, os classificadores são portadores de informação semântica, informação que pode ser expressa de diferentes modos, como formas fonologicamente dependentes (prefixos, sufixos, etc.) ou independentes (palavras sintaticamente individualizadas). Deste modo, um classificador é implementado como um operador semântico que possui expressão fonológica.

Portanto, a inclusão de componentes semânticos e fonológicos deve ser tratada como um fenômeno geral dos processos gramaticais, e, assim, implementados por parâmetros de núcleos. Temos, então, mais dois parâmetros para h : s , uma sequência de traços semânticos; e p , uma sequência de traços fonológicos. Como s e p devem ser visíveis ao operador característico k , e esses três operadores precisam “enxergar” o argumento interno i , a definição de h então resulta:

$$\begin{array}{ll}
 h = \lambda b k s p i e. & \\
 \lambda f. (f (k s p) i) e & \text{if } b i e \\
 error & \text{otherwise}
 \end{array}$$

A expressão acima já é verdadeiramente uma definição de núcleo. Ela mostra o operador característico do núcleo, k , processando um feixe de traços semânticos, s , e fonológicos, p ; a composição funcional $k s p$ é também um operador, que se aplica ao argumento interno, i . O destino natural desta composição $k s p i$ é aplicar-se ao argumento externo, e , tão logo o contexto externo f receba um valor apropriado. Observe que a função de seleção, b , não figura diretamente na estrutura do resultado bem-sucedido do núcleo, $\lambda f. (f (k s p) i) e$; ¹¹ b é um validador dos argumentos i e e .

Os quatro primeiros parâmetros — b , k , s e p — organizam a lógica interna do núcleo; por este motivo serão denominados *parâmetros interiores*. Os dois últimos parâmetros — i e e — são responsáveis pelo seu comportamento externo e serão denominados *parâmetros exteriores*.

Fica, assim, concluída a especificação de núcleo a ser utilizada no restante do trabalho. Todos os conceitos gramaticais a serem utilizados são obtidos por uma particular parametrização do operador h . Existe, portanto, um único núcleo, h , supostamente uma estrutura universal do sistema da linguagem, cujas parametrizações resultam em operadores especializados. Deste modo, a expressão “núcleo h ” é redundante, uma vez que não há outro núcleo; é suficiente mencionar *núcleo* ou h para fazer referência ao núcleo universal.

Um núcleo que apresente os parâmetros interiores preenchidos será denominado *núcleo parametrizado* e seu o resultado final computado tem a forma

$$\lambda i e. (\lambda f. (f ((k s p) i)) e),$$

em que somente os parâmetros exteriores permanecem disponíveis para o recebimento de valores e o núcleo encontra-se então preparado para participar de uma derivação sintática. Algumas vezes um núcleo parametrizado será chamado simplesmente de *núcleo*, por constituir um caso especial de h , mas deve-se ter em conta que existe um único núcleo, passível de diferentes parametrizações.

¹¹O resultado mal-sucedido é *error*.

As *categorias lexicais*, *nome* (N) e *verbo* (V), são exemplos de núcleos parametrizados. No modelo desta pesquisa, as *adposições* também são núcleos parametrizados, porém pertencem ao *sistema funcional* da GU, constituindo uma *categoria funcional*. O principal motivo de assim as considerar é a sua participação na formação de estruturas universais, como os verbos gramaticais *dar* e *ter*, tema ainda a ser abordado. Outra forte razão é o fato de as adposições comporem uma classe fechada de termos, qualquer que seja o idioma. Uma discussão sobre as adposições como termos funcionais é encontrada em Baker (2004, p. 303). Segundo a presente proposta, os *adjetivos* são caracterizados por determinadas configurações estruturais, não correspondendo propriamente a categorias lexicais ou funcionais.

5.4 Núcleos e traços semânticos e fonológicos

Supostamente, os núcleos parametrizados são parte da GU, mas o são de modo derivado, pois resultam da parametrização do núcleo universal por *operadores*, *feixes de traços semânticos* e *feixes de traços fonológicos*. Toda a pesquisa gira em volta desse processo, cujos elementos serão gradualmente introduzidos ao longo do trabalho.

Também supostamente, todos os operadores e parte dos traços semânticos pertencem à GU.

Serão tomados como pertencentes à GU os traços semânticos: *st*, estatividade; *ac*, atividade; *in*, incidência; *ec*, causalidade externa; *ic*, causalidade interna. Estes quatro traços são considerados operações cognitivas primitivas integrantes do sistema computacional da linguagem.¹² O traço *st* caracteriza aquilo que o falante compreende como estático (coisa ou estado), enquanto *ac* corresponde à interpretação dinâmica de algo (processo ou atividade); *in* indica a interação binária independente de ordem entre coisas e/ou processos; *ic* denota o surgimento espontâneo de algo estático ou dinâmico; *ec* designa o surgimento de algo (estático ou dinâmico) provocado, direta ou indiretamente, por outro algo.

Permanecerá em aberto a questão de se outros traços semânticos também pertencem à GU, ou se são específicos de determinadas línguas, mas realizados por mecanismos computacionais universais. Quase certamente, os traços indicadores de número (*pl*, *sing*, etc.) são universais. Há fundados questionamentos de se traços caracterizadores de gênero (*masc*, *fem* etc.) são traços semânticos, como os pertencentes à GU; seguramente não o são, como também não o são indicadores de coisas redondas ou compridas, etc., existentes em certas línguas. Comumente, características como essas são tratadas, em teoria gramatical, como *classificadores*. De todo modo, considerar-se-á que mesmo os classificadores portam alguma semântica e serão tratados como traços semânticos (embora não pertencentes à GU).¹³

¹²*Operações primitivas*: operações não definíveis segundo a teoria do sistema em que operam.

¹³Sob o ponto vista computacional, não há diferença entre os traços semânticos universais e não-universais, exceto que os últimos são adquiridos por aprendizado, diferentemente dos primeiros, que pertencem ao sistema linguístico biológico inicial de um falante.

Alguns traços semânticos não se expressam diretamente nas expressões pronunciadas, como o traço *st*, em português. Outros, produzem resultados fonológicos característicos, como o traço semântico *fem* (feminino) do português. Também os traços fonológicos produzem resultados, como se observa das expressões pronunciadas. Por este motivo, na presente proposta, os traços semânticos e fonológicos são operadores, uma vez que produzem resultados;¹⁴ não nos dedicaremos porém a explorar tais operadores. Para a pesquisa, eles são considerados operadores primitivos, não analisados. A única estrutura postulada para os traços semânticos e fonológicos é a de que eles se organizam em sequências, como $\langle sing, fem, st \rangle$ e *leved*. Em uma sequência, a ordem dos termos é importante.

Na teoria que estamos desenvolvendo, os traços semânticos são fundamentais às derivações: o modo como um constituinte se expressa fonologicamente depende da combinação de operadores e traços semânticos.

5.5 Exemplo de núcleo parametrizado

Pode-se, agora definir o nominalizador em *ura* como núcleo:

$$N_{URA} = h \beta n \langle sing, fem \rangle ur.$$

Uma vez que as quatro primeiras valências de *h* estão satisfeitas, permanecem em aberto apenas os argumentos interno e externo:

$$N_{URA} = \lambda ie. (\lambda f. (f ((n \langle sing, fem \rangle ur) i)) e).$$

Resulta, assim, um operador preparado para participar de uma derivação, à espera de dois valores. Forneçamos, para exemplificar, os valores para formar *levedura*. O núcleo N_{URA} irá então aplicar-se à (suposta) raiz *leved*. Dada a característica estática (ou não-dinâmica) desta raiz, postularemos que ela é dotada do traço semântico estativo,¹⁵ $\langle +st \rangle$,¹⁶ e obviamente do

¹⁴Segundo a proposta, o traço *fem*, que em português se expressa geralmente pela sufixação da fonologia *a*, não é um valor estático à disposição de um operador para se revelar. No contexto adequado, ele próprio é o operador que efetua a sufixação da fonologia. Do mesmo modo, a fonologia (i.e., o feixe de traços) *a* é um operador, produtor de resultados acústicos. Não há, assim, valores estáticos no modelo proposto; todos os elementos são operadores — alguns são operadores que se manifestam fonologicamente, outros são operadores “mudos”.

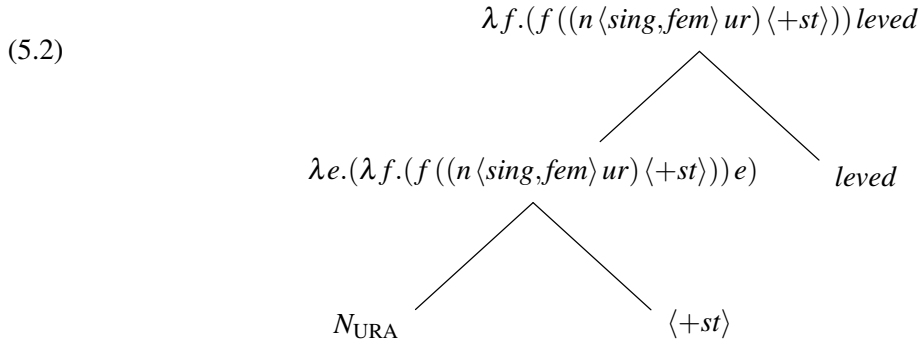
¹⁵A atribuição de um traço semântico a um item lexical básico é sempre uma postulação, que, entretanto, deve ser motivada; no caso, a motivação é dada pela característica estática da raiz. Não se trata de uma propriedade lógica ou matemática, mas de senso empírico.

¹⁶Outros nomes em *ura* podem ser derivados de raízes ativas (dotadas do traço de atividade *ac*), como *lavra-tura*. Tal fato significa que a interpretação estativa ou ativa dos nomes em *ura* decorre de propriedades lexicais da raiz em combinação com o nominalizador N_{URA} . Analisar as várias facetas do tema extrapola os limites desta pesquisa.

componente fonológico *leved*. Portanto:

$$N_{\text{URA}} \langle +st \rangle \text{leved} \mapsto \lambda f. (f ((n \langle sing, fem \rangle ur) \langle +st \rangle)) \text{leved}.$$

A derivação pode ser ilustrada por meio de um diagrama.



O topo do diagrama mostra o *nP* a ser pronunciado como 'levedura'.¹⁷

A estrutura do constituinte codifica, explicitamente, a sua história derivacional. No exemplo, a estrutura resultante é mais simples do que à primeira vista parece. Se considerarmos que o operador composto $n \langle sing, fem \rangle ur$ atua como uma unidade funcional e denotarmos tal unidade por n_{URA} ,¹⁸

$$n_{\text{URA}} = n \langle sing, fem \rangle ur,$$

a estrutura do *nP* derivado pode ser equivalentemente representada por

$$\lambda f. (f (n_{\text{URA}} \langle +st \rangle)) \text{leved}.$$

Isto é: o constituinte é, canonicamente, composto por um operador (n_{URA}) e seus dois argumentos ($\langle +st \rangle$ e *leved*).

Uma estrutura pode ser apresentada em diferentes graus de abstração equivalentes entre si. O constituinte do topo do diagrama mostra a que pode ser chamada de “representação de baixo nível”, em que os elementos computacionais aparecem de modo explícito. Este estilo de representação facilita o acompanhamento dos resultados das computações, passo a passo. Já no estilo de representação acima, algumas estruturas recorrentes aparecem representadas como um todo, o que as torna mais amenas à análise estrutural dos constituintes; as relações estruturais são mais facilmente visíveis. Pode-se dizer que são “representações de alto nível”,

¹⁷Na **Parte III Estruturas Verbais**, analisaremos como se forma a expressão fonológica de um constituinte.

¹⁸A substituição de um operador composto e empregado de modo sistemático, como $n \langle sing, fem \rangle ur$, por uma definição, como n_{URA} , é uma instância do processo de *lexicalização*. O operador (no caso, n_{URA}) é aprendido pelo falante e registrado em seu léxico; i.e., a combinação específica $n \langle sing, fem \rangle ur$, computacionalmente provida pela GU, é lexicalizada, encapsulada em um termo de mais alto nível, uma espécie de sinônimo, utilizada como uma unidade pelo falante.

por favorecerem as análises mais abstratas, de certo modo mais afastadas da execução dos processos. Sempre que possível, dar-se-á preferência às representações de mais alto nível, por facilitarem as análises estruturais; de todo modo, as computações estão inteiramente embutidas nas representações mais abstratas, podendo ser abertas sempre que necessário. Não há diferença computacional entre os níveis de abstração.

5.6 O operador característico

Até o momento, temos uma especificação de núcleo de grande generalidade, que parece suficiente à solução das estruturas analisadas no restante do trabalho. Entretanto, ainda há uma questão, de certo modo velada, que precisa ser resolvida de modo sistemático, sob a perspectiva de uma estrutura universal de núcleo.

Chegamos ao ponto de maior complexidade da teoria do núcleo.

Adotamos a hipótese de trabalho de que os núcleos pertencem à GU, embora cada língua selecione uma parte deles para compor o seu sistema computacional. Dada a sua característica universal, cada núcleo é estruturalmente único em todas as línguas. Entretanto a estrutura única de um determinado núcleo pode se expressar de diferentes modos nas línguas particulares.

A estrutura genitiva, por exemplo, seria codificada por um núcleo, sendo portanto pertencente à GU. No caso do inglês, a estrutura genitiva se manifesta de, pelo menos, dois modos: 'Book of Mary', por uma preposição; 'Mary's book', por uma desinência. Temos, assim, um fenômeno que merece atenção: no mesmo idioma, um mesmo núcleo se manifestando de diferentes formas.

Efetivamente, nada há de estranho no fato de um núcleo se manifestar de modos diferenciados em um mesmo idioma. Todo o sistema verbal de qualquer língua se fundamenta nesta característica dos núcleos, como o evidenciam as diferentes expressões de tempo e modo. A questão é que um núcleo pode conter opções de expressão, mas não pode ele mesmo selecionar uma das opções. A escolha da opção é uma informação ainda não disponível no momento em que o núcleo entra em uma derivação. Por conseguinte, a escolha, ou a determinação, do modo como um núcleo irá se expressar cabe a um outro núcleo, incluído após ele na derivação.

Isto significa que o núcleo que se expressa pela relação genitiva deve permanecer visível a outros núcleos, ao longo de uma derivação.

No caso de 'Book of Mary', a visibilidade é imediata, manifestada pela preposição. A preposição exhibe a individualidade sintática do núcleo de um modo seguramente bastante claro. Já em 'Mary's book', o núcleo parece ter sido absorvido por um *nP*, *Mary*, perdendo a sua individualidade. Contudo, como o núcleo em questão é único, do ponto de vista estrutural, não pode ocorrer que ora ele tenha visibilidade sintática, ora não a tenha; assim, independentemente do modo de se expressar, ele deve permanecer visível a outros núcleos ao longo da derivação.

A preservação da individualidade sintática dos núcleos não é exclusividade do genitivo

inglês. A diferença entre 'Maria vai falar' e 'Maria falará', como também a similaridade entre 'João tem um livro' e 'Maria deu um livro a João',¹⁹ reside justamente nos diferentes modos com que determinadas combinações de núcleos podem se expressar, fato que subentende a visibilidade sintática dos núcleos envolvidos na formação das expressões. Esta é uma das propriedades mais fundamentais da gramática, a expressão contextual dos núcleos, o que pressupõe a sua individualidade e visibilidade estrutural.

Analisemos, agora, a questão com relação ao núcleo h , tomando como exemplo o pP , anteriormente derivado, $\lambda f.(f(p\text{Maria}))\text{casa}$. Do modo como se encontra definido, nada pode impedir que o operador p se aplique ao nP , *Maria*, produzindo um certo resultado. Tem-se então que o operador p é substituído pelo resultado da computação $p\text{Maria}$. Ou seja, o operador p perde a sua visibilidade estrutural, somente se manifestando por meio de uma nova expressão, o resultado de pnP .

A questão é absolutamente geral: o resultado da aplicação bem-sucedida de um núcleo possui a estrutura

$$\lambda f.(f((ksp)i))e.$$

Isto significa que o operador característico k , aplicado a uma semântica s e a uma fonologia p , constitui um novo operador composto ksp . Este operador aplica-se obrigatoriamente ao argumento interno, i . Ou seja, o que se torna visível na estrutura é a fusão do operador ksp com o argumento interno. Assim, o operador k , que exprime a lógica de um certo núcleo x , caracterizador do constituinte como um xP de determinado tipo, perde a sua identidade sintática: ele se funde com argumento interno. Este fato é incompatível com uma propriedade fundamental da gramática, a de que a expressão fonológica de um constituinte resulta da composição de seus núcleos.

Estamos diante do fato de que o operador característico k também possui uma estrutura que precisa ser investigada. Para fazê-lo, em primeiro lugar é preciso compreender o que significa a aplicação ksp .

Para a formação de nomes, k recebe o valor n , o operador nominalizador, particular de uma determinada língua; analogamente, para a formação de verbos, k recebe v ; para preposições, p . Além de categorizar nomes e verbos, e de relacionar termos por meio de preposições, veremos, na continuação do texto, que os núcleos também são responsáveis pela inclusão, nas derivações, de tempo, aspecto, etc.. esses núcleos também incluem seus operadores, *tense*, *asp*, etc.. Deste modo, k denota um operador categorizador ou modalizador; k contém o tratamento que o sistema de uma língua particular especifica para uma determinada categoria gramatical.

Evidentemente, a investigação da lógica interna desses operadores vai além dos objetivos do trabalho. Analisar o sistema de formação de nomes ou verbos de uma língua constitui, por si, uma pesquisa. O objeto da minha investigação é a máquina computacional em que este sistema funciona. Assim, a análise tem em vista o relacionamento dos os núcleos entre si e com

¹⁹Como veremos em 8.8.3, p. 137, os verbos *ter* e *dar* constituem diferentes expressões de uma cadeia de núcleos.

seus argumentos.

Todavia, como vimos no caso dos nomes em *ura*, o núcleo N pode ser especializado para realizar certas formações particulares. Daí surgiram especializações como $N \langle sing, fem \rangle ur$, em que um núcleo aparece parametrizado; e também as correspondentes especializações dos operadores, $n \langle sing, fem \rangle ur$.

Assim, o operador n realiza o tratamento geral dos nomes de uma língua, em termos de suas categorias e recursos formais, como gênero, número, sufixos, prefixos, etc..

O ponto a ser destacado é que uma aplicação como $n \langle sing, fem \rangle ur$ é também um operador; um operador que parametriza a lógica de n , internalizando uma certa semântica e fonologia.

Ocorre, entretanto, que a semântica interna de um núcleo (e isto inclui a semântica com que ele foi parametrizado) possui implicações sintáticas. (Também a fonologia do núcleo apresenta potencialmente implicações sintáticas, fato que será deixado à parte.)

Por exemplo, como veremos, o núcleo verbalizador V pode ser parametrizado por uma semântica ativa, ac , e tal fato implica que o núcleo então formará verbos que exibem um certo caráter de atividade. A parametrização $V ac$ será definida como o núcleo V_{DO} , o “verbalizador ativo”, embora, em alguns casos, ele verbalize estruturas estativas. A compatibilidade de V_{DO} com estruturas estativas depende inteiramente de traços semânticos do núcleo e de seus parâmetros.

Assim, não é apenas a visibilidade sintática do núcleo que deve ser preservada na estrutura, mas sim a visibilidade do núcleo combinado com seus argumentos semânticos e fonológicos, pois é esta composição que determina o comportamento morfo-sintático dos categorizadores e modalizadores.

Portanto, quando se afirma que um operador característico k deve preservar a sua individualidade sintática estrutural, enuncia-se uma verdade incompleta: verdadeiramente, o que deve permanecer visível em um constituinte é o operador composto $k s p$.

5.6.1 Estrutura interna do operador característico

Podemos, assim, localizar o problema a ser resolvido, quando se fala em preservar a integridade sintática de um núcleo. Quanto à estrutura de um constituinte, $\lambda f.(f((k s p) i)) e$, o operador composto $k s p$ deve ser impedido de se aplicar ao argumento interno, i . Portanto, todo o constituinte precisa ser mantido em estado de processamento suspenso, *exceto* o operador composto $k s p$, a lógica (interna) do núcleo. Quanto ao relacionamento externo, o operador $k s p$ deve ser capaz de receber informação originada de outros núcleos, incluídos após ele na derivação; isto é, deve ser sensível a um contexto externo, fonte da informação que será “infiltrada” no constituinte por meio da variável f .

A chave da solução está em se definir a estrutura interna de k de modo a se obterem os efeitos citados acima:

- (a) Para ser sensível a um contexto externo, o operador k deve ter o formato:

$$k = \lambda c. c w,$$

em que c é o contexto externo e w é uma operação.

- (b) Para apresentar aplicação postergada ao argumento interno, k deve ter o formato:

$$k = \lambda ic. (c w) i.$$

- (c) Não podemos nos esquecer que o operador k também “enxerga” a semântica, s , e a fonologia, p , específicas do núcleo, assim:

$$k = \lambda spic. (c (w s p)) i.$$

A expressão mostra a *estrutura interna* do operador característico k : k contém um processo w que depende de um contexto externo c para se aplicar ao argumento interno i , sendo que w processa uma semântica s e uma fonologia p .

Levando-se em conta a estrutura interna do operador característico, todo constituinte pode ser representado diretamente com o operador k ou com k substituído por sua estrutura interna, como é mostrado pelas expressões equivalentes mostradas abaixo:

$$\lambda f. (f (k s p) i) e,$$

e

$$\lambda f. (f (\lambda c. (c (w s p)) i)) e.$$

Os diferentes modos de representação de um constituinte atendem a finalidades variadas quando da análise de uma derivação, tema analisado na p. 49, ao se considerarem os graus de abstração de uma representação.

5.6.2 Exemplo de estrutura interna

Para exemplificar, reanalisemos o nominalizador em *ura* tornando explícita a estrutura interna do operador nominalizador n .

Como vimos na pag. 49, a derivação do nome pronunciado como *levedura* gerou o constituinte

$$\lambda f. (f (n_{\text{URA}} \langle +st \rangle)) leved.$$

O constituinte mostra o operador n_{URA} definido como

$$n_{\text{URA}} = n \langle \text{sing}, \text{fem} \rangle \text{ur}.$$

As expressões acima continuam válidas, porém agora estamos munidos de uma definição do operador n que impede a aplicação de n_{URA} ao argumento interno $\langle +st \rangle$, aplicação que, caso ocorresse, implicaria a falta de visibilidade estrutural de n_{URA} , segundo as análises anteriores.

Vejamos como a integridade de n_{URA} efetivamente fica preservada na derivação:

- (a) Em primeiro lugar, devemos observar que n possui a estrutura interna

$$n = \lambda \text{spic}.(c(wsp))i,$$

a estrutura de um operador gramatical, segundo as análises anteriores.

- (b) Aplicando-se n , como definido acima, a $\langle \text{sing}, \text{fem} \rangle$ e ur , a aplicação $n \langle \text{sing}, \text{fem} \rangle \text{ur}$, da definição de n_{URA} , preenche as duas primeiras valências de n , respectivamente s e p .

- (c) Deste modo, n_{URA} , em notação de baixo nível, resulta:

$$n_{\text{URA}} = \lambda ic.(c(w \langle \text{sing}, \text{fem} \rangle \text{ur}))i,$$

um operador “à espera” de aplicar-se ao argumento (interno) i .

- (d) Não devemos nos esquecer que o constituinte que deriva ‘levedura’ possui a estrutura

$$\lambda f.(f(n_{\text{URA}} \langle +st \rangle)) \text{leved},$$

em que $\langle +st \rangle$ é o argumento interno.

- (e) Agora, substituamos n_{URA} , do constituinte acima, pela sua expressão de baixo nível. Neste caso, o argumento interno do constituinte, $\langle +st \rangle$, satura a primeira valência em aberto, λi , de n_{URA} .

- (f) Resulta, então, a estrutura do constituinte que deriva ‘levedura’, com todos os elementos computacionais exibidos de modo explícito:

$$\lambda f.(f(\lambda c.(c(w \langle \text{sing}, \text{fem} \rangle \text{ur})) \langle +st \rangle)) \text{leved}.$$

O constituinte recém-derivado, deve-se reiterar, é equivalente ao de mais alto nível

$$\lambda f.(f(n_{\text{URA}} \langle +st \rangle)) \text{leved},$$

diferindo no nível da notação.

5.7 O paradigma gramatical

Temos, no exemplo anterior, a derivação de um constituinte em que se utilizou a estrutura do operador característico de modo explícito, no caso, o operador do núcleo nominalizador, n .

Entretanto, um ponto foi passado por alto, por simplicidade de análise. Ao se investigar a estrutura interna de n , nada se disse a respeito de w , tomado simplesmente como “um operador”.

Voltando à especificação de n , $\lambda spic.(c(wsp))i$, verificamos que n é parametrizado, nesta ordem, por uma *semântica*, uma *fonologia*, um *argumento interno* e um *contexto externo* — termos que são recebidos “de fora”. Entretanto, de onde veio w ?

O operador n é um categorizador nominal e a especificação de n , na realidade, estabelece o cenário para que a categorização possa acontecer. Ao ocorrer uma nominalização, o valor de todas as variáveis presentes no constituinte nominalizado pode mudar, exceto o valor de w . Ou seja, w é introduzido por n , não é recebido do ambiente externo ao constituinte; por este motivo, w não figura como um dos parâmetros de n na lista de parâmetros $\lambda spic$.

Esta é a função de um operador característico k : criar o ambiente para que um processo gramatical w possa ocorrer.

Cada operador característico introduz, então, o seu w particular. Para deixar patente este fato, convencionaremos que o w específico de n é denotado por w_n . Do mesmo modo, há também os operadores w_v , w_p etc., para a formação de verbos, preposições, etc.. A convenção tem também a vantagem de não precisarmos ser muito criativos para dar nomes a tais operadores.²⁰

Deste modo, a especificação formalmente rigorosa de um operador característico é

$$k = \lambda spic.(c(w_ksp))i.$$

No caso da nominalização, o operador w_n contém o processo específico de nominalização de uma dada língua. Por este motivo, chamaremos o operador w_k de *paradigma gramatical* da categoria k , ou, simplesmente, “paradigma de k ”. Nesta convenção, o termo “paradigma” é usado com o sentido de “um processo padrão”, e não de “um rol de possibilidades”.

Um ponto importante é o de cada falante possuir os seus operadores w_k particulares. Seguramente o meu operador w_n é diferente do seu operador w_n , as possíveis diferenças sendo balizadas pelas funções de seleção dos núcleos envolvidos nas derivações. Assim, os paradigmas gramaticais, w_k , codificam a língua de cada falante, inseridos em ambientes sintáticos que

²⁰Há também outro modo de se compreender w_k : w é o processo geral que especifica todas as categorizações gramaticais de uma dada língua e k é um parâmetro que seleciona uma parte deste processo. Assim, w_k é um operador de seleção de uma lógica gramatical particular k , situada em um ambiente maior, w . Sob esta interpretação, cada operador característico parametriza w de um modo especial, produzindo como resultado uma categorização gramatical particular, como w_n (o operador lexical formador de nomes) e w_p (o operador funcional formador de preposições).

são os mesmos para todos os seres humanos.

Os ambientes sintáticos são relativamente muito poucos, todos definidos por parametrizações do núcleo h ; e todos são variações de uma única estrutura: $\lambda f.(f((ksp)i))e$. Isto é, todo ambiente sintático possui a estrutura de um operador ksp , atuante sobre dois termos, i e e , sob um contexto não especificado f , segundo a estrutura de um *pair*.

5.8 Paradigma gramatical e argumentos do núcleo

Voltemos agora ao núcleo h , do qual toda computação bem-sucedida produz o resultado

$$\lambda f.(f((ksp)i))e,$$

em que i e e , como sabemos, são os argumentos interno e externo do núcleo.

Considerada a estrutura do constituinte acima, a composição do contexto f com o operador ksp possui i e e como argumentos interno e externo, respectivamente. Tal fato significa que o operador característico k possui o argumento externo em seu domínio de funcionalidade: existe um operador $f((ksp)i)$ e este operador aplica-se a e . Portanto, o argumento externo se situa em posição estrutural para ser afetado pelo operador característico k . Tal fato, em si, nada tem de notável, uma vez que e é o argumento de um núcleo e k é o operador que realiza a funcionalidade do mesmo núcleo.

Porém, consideremos a estrutura interna de k ,

$$\lambda spic.(c(w_ksp))i,$$

e levemos em conta que a única ação de k é criar o ambiente para a inserção do paradigma gramatical w_k .

Temos agora um fato interessante e teoricamente muito importante: o paradigma gramatical w_k não possui posição estrutural para ter e como um de seus argumentos.

Se desconsiderarmos, por um momento, o contexto c , a composição w_ksp possui p como argumento interno e i como argumento externo. Isto significa que, isolado do seu contexto, o paradigma gramatical não possui estrutura suficiente para atuar sobre o argumento externo do núcleo. Dito de modo simples: w_k não “enxerga” e .

Se consideramos que, por hipótese, todo operador é absolutamente agnóstico quanto aos contextos em que pode ser inserido, teremos que w_k , em sua lógica interna, não pode conter especificações a respeito do argumento externo do núcleo, e , uma vez que qualquer especificação dependeria da colaboração com o contexto c , do qual w_k não é ciente.

Isto explica a enorme dificuldade que um núcleo possui para atuar sobre o seu argumento externo: o argumento externo, embora seja argumento do núcleo, não é argumento do seu paradigma gramatical.

Um núcleo só pode atuar sobre o seu argumento externo em colaboração com outro núcleo; não pode contar com a sua própria lógica gramatical para isso.

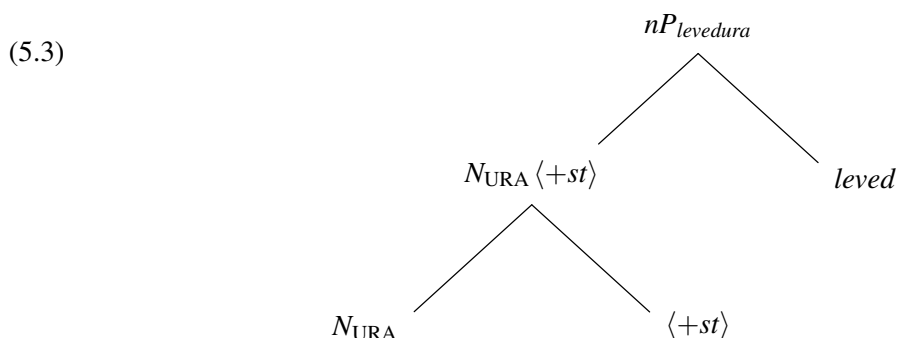
Contudo, parece haver uma família linguística em que ocorre o fenômeno da marcação do argumento externo pelo núcleo. A estrutura genitiva, em Kamaiurá, se caracteriza pela prefixação do fonema denotado por *r* ao argumento externo. Temos um exemplo em Seki (2000, p. 300).

- (23) wyrapiã rupi'a
 wyrapiã *r-* upi'a
 passarinho Rel- ovo
 'ovo de passarinho'

A formação do genitivo mediante a prefixação do argumento externo é o processo comumente utilizado pela família linguística tupi. Se as análises acima estiverem corretas, o genitivo tupi incorpora a ação de algum núcleo externo, segundo um processo bastante raro, considerado o custo computacional, mas que não é teoricamente impossível. Tudo considerado, a estrutura genitiva da família tupi pode mostrar-se uma contribuição das línguas sul-americanas à teoria da linguagem.

5.9 Núcleo e diagramas

Uma parte da notação utilizada deixa subentendida a estrutura funcional e semântica dos núcleos. Pode-se desenhar um diagrama como o abaixo, em que a complexidade computacional se encontra implícita,



como também é possível exibir toda a estrutura computacional dos constituintes participantes da derivação:

(5.4)

$$\begin{array}{c}
 \lambda f.(f((n \langle sing, fem \rangle ur) \langle +st \rangle)) leved \\
 \swarrow \quad \searrow \\
 \lambda e.(\lambda f.(f((n \langle sing, fem \rangle ur) \langle +st \rangle)) e) \quad leved \\
 \swarrow \quad \searrow \\
 N_{URA} \quad \langle +st \rangle
 \end{array}$$

Independentemente do grau de abstração em que um constituinte é representado, nada há nos diagramas que seja apenas uma notação a ser complementada com explicações; e nenhum papel desempenham os diagramas nas derivações, exceto a facilidade de visualização.

5.10 Núcleo e categorias gramaticais

Determinadas atribuições de valores aos parâmetros interiores apresentam como resultado a especificação das categorias gramaticais. Por exemplo, a categoria (lexical) dos verbos é definida pela parametrização

$$V = h\beta v,$$

e o verbalizador estativo, V_{BE} , é definido pela parametrização

$$V_{BE} = V \langle +st \rangle null = h\beta v \langle +st \rangle null.$$

Não nos deteremos, neste momento, na explicação dos valores dos parâmetros, o que será feito na continuação do texto. Muitos outros núcleos serão especificados e todas as derivações são efetuadas pela aplicação de núcleos a argumentos.

Um operador como V_{BE} , cujos parâmetros interiores se encontram saturados, possui em aberto apenas as valências dos argumentos exteriores, i e e . Portanto, V_{BE} pode ser expresso ainda de um outro modo:

$$V_{BE} = \lambda ie.(\lambda f.(f((v \langle +st \rangle null) i)) e).$$

Observe-se que a expressão V_{BE} é muito mais amena aos olhos do que o seu equivalente $\lambda ie.(\lambda f.(f((v \langle +st \rangle null) i)) e)$.

Operadores de alto nível, como V_{BE} e o já visto N_{URA} , deixam subentendida a complexidade da estrutura interna do núcleo e possuem em aberto apenas os argumentos exteriores, sendo, assim, preferencialmente usados em derivações.

As categorias gramaticais são núcleos parametrizados (conceitos complexos) encapsu-

lados em definições. Expressões como N , V , V_{DO} , e outras a serem analisadas, são centros de funcionalidade e semântica operacional, máquinas conceituais que atuam segundo uma lógica interna.

5.11 Definição formal com assinaturas funcionais

A definição de núcleo até aqui desenvolvida está operacionalmente completa, para os objetivos da pesquisa, mas não inclui a assinatura dos operadores. A assinatura dos operadores atende às finalidades mais formais da proposta, mas pode ser passada por alto em um primeiro contato com o tema. Como resumo, e para apresentar a definição formalmente explícita de núcleo, incluo a formalização do operador h na definição a abaixo.

Definição 5. Um núcleo é um operador h , formador de expressões linguísticas, dotado da seguinte estrutura:

$$(i) \quad \begin{aligned} h: B \rightarrow K \rightarrow \langle S \rangle \rightarrow \langle P \rangle \rightarrow T \rightarrow T \rightarrow pair\ T\ T \\ h = \lambda b k s p i e. \quad \begin{array}{ll} \lambda f. (f\ k\ s\ p\ i)\ e & \text{if } b\ i\ e \\ error & \text{otherwise} \end{array} \end{aligned}$$

(ii) Em que os parâmetros têm os seguintes significados:

- a. b é um critério de seleção aplicável aos valores dos parâmetros i e e , respectivamente o primeiro e o segundo parâmetros exteriores do núcleo; ou seja, b é uma função booleana que possui a assinatura $b: T \rightarrow T \rightarrow Bool$;
- b. k é uma operação aplicável aos componentes semântico, s , fonológico, p , e ao argumento interno i ; k introduz o paradigma morfo-sintático do núcleo, w_k , sendo parametrizada por uma operação κ de estrutura $\kappa = \lambda s p i. \lambda c. (c\ (w_k\ s\ p))\ i$, em que c é um contexto externo e w_k possui a assinatura

$$w_k: \langle S \rangle \rightarrow \langle P \rangle \rightarrow T \rightarrow T;$$

deste modo, k possui a assinatura

$$k: \langle S \rangle \rightarrow \langle P \rangle \rightarrow T \rightarrow pair\ (W\ S\ P)\ T,$$

em que W é o tipo do paradigma w_k .

- c. s é um feixe de traços semânticos, $\langle \sigma_i \rangle$, o acréscimo semântico específico do núcleo;
- d. p é um feixe de traços fonológicos, $\langle \phi_i \rangle$, o acréscimo fonológico específico do núcleo;

- e. i é o primeiro parâmetro exterior do núcleo; ou seja, o valor atribuído a i é o argumento interno do núcleo;
- f. e é o segundo parâmetro exterior do núcleo; ou seja, o valor atribuído a e é o argumento externo do núcleo.

5.12 Propriedades estruturais do contexto

O contexto externo c possui importantes propriedades estruturais que determinam os resultados de todo o restante do trabalho.

Com uma parametrização bem-sucedida (aquela que não é abortada pela função booleana β nem pela seleção argumental do operador κ), um núcleo h produz como resultado o constituinte

$$\lambda f.(f(\kappa\langle\sigma_i\rangle str i))e.^{21}$$

Tendo em vista que a operação κ possui a estrutura

$$\kappa = \lambda spi.\lambda c.(c(w_k s p))i,$$

o resultado final da parametrização de h é um núcleo parametrizado de estrutura

$$\lambda f.(f(\lambda c.(c(w_k\langle\sigma_i\rangle str))i))e.$$

Um exemplo de núcleo parametrizado, ilustrativo da estrutura acima, é o já analisado N_{URA} ,²²

$$\lambda f.(f(\lambda c.(c(w_n\langle sing, fem\rangle ur))i))e,$$

em que: $w_k = w_n$, $\langle\sigma_i\rangle = \langle sing, fem\rangle$ e $str = ur$.

Na seção 5.8, p. 56, já analisamos a propriedade fundamental do contexto c , que é a de postergar a aplicação do operador composto $w_k\langle\sigma_i\rangle str$ ao argumento interno i . Deste modo, o operador do núcleo e o argumento interno preservam a sua individualidade sintática ao longo da derivação, permanecendo disponíveis para operações futuras. Por exemplo, o argumento interno mantém-se apto a receber caso e o operador do núcleo preserva a capacidade de participar de operações de concordância e de expressão de aspecto, apenas para citar algumas possíveis operações.

A propriedade que acabamos de analisar é de natureza puramente computacional, resultante do conceito de aplicação funcional, em conjunto com a noção de *pair*, uma vez que o contexto c estabelece uma relação de *pair* entre o operador do núcleo e i .

²¹Na realidade, o constituinte possui a estrutura $\lambda ie.(\lambda f.(f(\kappa\langle\sigma_i\rangle str i))e)$. Para facilitar a exposição, nas análises desta seção foi omitida a estrutura de parâmetros λie dos constituintes.

²²Análises na p. 54, em que o paradigma gramatical é utilizado em um formato simplificado, w .

Há, porém, uma outra propriedade de c , de natureza especificamente linguística e muito profunda, que possui implicações em toda a teoria linguística da pesquisa.

Analisemos a estrutura de w_k , o paradigma morfo-sintático do núcleo. O paradigma w_k é parametrizado por traços semânticos, s , e fonológicos, p ; este operador composto, $w_k s p$, aplica-se a um termo, t , produzindo como resultado um termo t' , *do mesmo tipo de t* . Se t é uma raiz, então t' é uma raiz; se t é um νP , então t' é um νP , e assim por diante.

Portanto, a estrutura do paradigma gramatical w_k contém suposições extremamente fortes: (i) w_k é necessariamente parametrizado por traços semânticos e fonológicos, quaisquer que sejam os tipos dos argumentos interno e externo do núcleo; (ii) ademais, o resultado da computação iniciada por w_k é necessariamente do mesmo tipo do argumento interno, não importa qual seja o tipo deste argumento. Tais suposições encontram-se codificadas na assinatura de w_k , $w_k: \langle S \rangle \rightarrow \langle P \rangle \rightarrow T \rightarrow T$, a qual restringe as configurações de argumentos possíveis para este operador.

Retornemos à estrutura do núcleo h , adequadamente parametrizado:

$$\lambda f.(f(\kappa \langle \sigma_i \rangle stri))e.$$

A estrutura deixa claro que toda a computação do núcleo tem início no operador κ . Mas, pela estrutura interna de κ , a sua computação interna tem início em w_k . Isto significa que a computação do núcleo h *sempre* atingirá um resultado do tipo

$$\lambda c.(c(w_k s p))i.$$

Caso não haja participação do contexto c , a computação interna do núcleo, iniciada em w_k , terminará necessariamente computando um resultado i' , do mesmo tipo do argumento interno i , segundo a especificação da assinatura de w_k ; este resultado nunca será a aplicação de i' a e , situação em que i' , o resultado da computação do paradigma do núcleo, processa e , o argumento externo. Está implícito neste fato que o operador gramatical κ *não contém prescrições a respeito do argumento externo*. Temos assim a fundamentação computacional das propriedades do paradigma gramatical já analisadas na p. 56 de modo mais intuitivo: *e, o argumento externo do núcleo h , não é diretamente “enxergado” pela estrutura interna do operador gramatical κ* . Ou seja, o contexto c é necessário à computação do argumento externo pelo paradigma gramatical do núcleo.

O paradigma w_k depende apenas da parametrização por uma semântica s e uma fonologia p para atuar sobre o argumento interno i , não sendo para tanto necessário o contexto c . Tal fato explica, por exemplo, a capacidade de uma preposição atribuir caso ao seu argumento interno, independentemente do ambiente sintático em que está incluída.

Entretanto, é preciso ter em conta que o argumento interno i se encontra afeito a duas relações estruturais: (a) a relação iniciada pelo paradigma w_k , $(w_k s p)i$; e (b) a relação iniciada pelo contexto c , $(c(w_k s p))i$. Cada uma dessas relações introduz diferentes efeitos sobre i ,

como ocorre, por exemplo, na associação *queimar campo*, em que há um paradigma verbal, w_v , atuando sobre um argumento interno nominal, o nP *campo*. A relação estrutural (a) caracteriza o nP como *tema* de w_v ; já a relação estrutural (b), introduzida pelo contexto externo, caracteriza o nP como *sujeito* (de um verbo inacusativo, ‘O campo queimou’) ou como *objeto* (de um verbo transitivo, ‘A geada queimou o campo’), entre outras possibilidades, todas dependentes do contexto c .

Deve-se, contudo, observar que a computação contextual (caracterização “sintática” do nP como sujeito ou objeto; atribuição de caso) não altera a relação estrutural do argumento interno com o paradigma do núcleo (caracterização “semântico-temática” do nP), responsável pela semântica fundamental do constituinte. Este é o resultado esperado, tendo em vista que a computação do argumento interno i pelo paradigma w_k independe de c , como visto anteriormente.

Baseado no fato empiricamente observado de uma preposição nunca concordar com o seu argumento externo, Baker conclui que as preposições são predicados monoargumentais: “*There [(Baker 1996, ch. 9, pp. 400–404)] I pointed out that PPs never agree with their putative subjects, even in heavily head-marking languages like Mohawk*”, Baker (2004, p. 315).²³ Com base nas análises precedentes, sabemos que nenhum núcleo é capaz de estabelecer concordância com o seu argumento externo, apenas com o seu argumento interno (fato encontrado por Baker); a concordância de um núcleo com o seu argumento externo é sempre disparada por contextos externos ao núcleo. Assim, o argumento principal de Baker não é suficiente para caracterizar as preposições como predicados monoargumentais. A alegada ausência de concordância de uma preposição com o seu argumento externo, diferentemente ao que ocorre com relação aos verbos, se deve aos diferentes contextos em que preposições e verbos encontram-se incluídos, fato alheio à estrutura argumental das duas categorias.

Podemos agora revisitar uma importante propriedade anteriormente encontrada (p. 56): o argumento externo é argumento do núcleo, mas não do paradigma gramatical do mesmo núcleo; o núcleo, h , pode atuar sobre o seu argumento externo, mas não o faz por meio do seu operador gramatical k . O núcleo somente atua sobre o seu argumento externo por meio de “material” fornecido por outros núcleos, através do contexto c , um ponto de passagem. Este fato é implicado pelo modelo computacional.

Como consequência, o “acabamento” que um constituinte precisa ter para ser pronunciado (isto é, a saturação da posição argumental f), no caso do argumento externo, deve ser proporcionado por um núcleo externo ao constituinte.

O trabalho propriamente linguístico necessário para definir um núcleo reside na seleção do paradigma w_k , a única operação específica do núcleo, como é mostrado nas análises anteriores. Temos, então, um importante guia prático para a análise interna de w_k , o paradigma morfo-sintático do núcleo: a lógica deste paradigma somente contempla o argumento interno; é

²³“Lá eu assinaiei que PPs nunca concordam com seus supostos sujeitos, mesmo em línguas fortemente marcadoras de núcleos como o Mohawk ...”

inútil tentar resolver problemas que envolvem a morfo-sintaxe do argumento externo ao analisar a morfo-sintaxe própria do núcleo.

O paradigma w_k do núcleo pode instalar no argumento interno i operações que atuem sobre o argumento externo e (como o traço semântico *fem* introduzido por N_{URA} , que produz consequências fonológicas sobre o argumento externo), entretanto a aplicação dessas operações a e somente ocorrerá em virtude de operações provenientes de núcleos externos, por meio do contexto c — isto é o que mostram as análises precedentes. O resultado final da computação do núcleo será $(w_k s p) i$ processando e , *segundo as regras estabelecidas pelo contexto c a partir de informação de origem externa ao núcleo.*

Há uma consequência de muitas implicações a respeito do paradigma w_k : uma vez que (a) a computação de $(w_k s p) i$ sempre resulta em um constituinte i' *de mesmo tipo* do argumento interno i , e que (b) a computação final do núcleo será $i' e$, resulta que (c) *não existe a possibilidade de o paradigma do núcleo determinar a ordem segundo a qual os argumentos interno e externo serão pronunciados*, exceto se a ordem for determinada pelo contexto c . Esta é uma consequência de o operador composto $(w_k s p) i$ não “enxergar” e . Ou seja, *a ordem de pronuncia dos argumentos de um núcleo é sempre externa ao paradigma do núcleo.*

Levando-se em conta que os dois argumentos, i e e , são ambos processados pelo núcleo (um truismo), e que o núcleo não possui funcionalidade suficiente para determinar a ordem em que os dois constituintes serão pronunciados, tem-se por conseguinte que a informação mínima especificada para o contexto c por um núcleo externo é a definição da ordem em que i e e serão pronunciados.

Uma consequência prática deste fato é a de que é inútil buscar dentro de um núcleo uma lógica que determine a ordem em que os argumentos deste núcleo são pronunciados; esta lógica estará necessariamente a cargo de um núcleo externo. Tal fato resulta do modelo computacional proposto, sem estipulações, e é compatível com um dos pilares da Teoria dos Princípios e Parâmetros, que estabelece a ordem de pronúncia dos constituintes sentenciais como o valor de um parâmetro externo à sentença.

Parte IV

Derivações sintáticas

Capítulo 6

Estruturas básicas

6.1 Tipos de núcleos

Vimos anteriormente que toda expressão linguística é o resultado de uma parametrização particular do núcleo, o operador h . Relembremos que h possui a especificação

$$h = \lambda b k s p i e . (\lambda f . (f ((k s p) i)) e \dots),$$

em que:

- (a) b é a função booleana responsável pelas restrições seletivas do núcleo, definidora das características dos argumentos interno e externo a que o núcleo pode se aplicar de modo bem sucedido;
- (b) k é o operador portador da lógica específica do núcleo; aquilo que define o núcleo como uma preposição, um categorizador de nomes, de verbos, etc.;
- (c) s é a semântica, $\langle \sigma_i \rangle$, particular do núcleo; aquilo que, por exemplo, como veremos nos próximos capítulos, diferencia o verbalizador estativo V_{BE} , dotado da semântica $\langle +st \rangle$, do verbalizador ativo V_{DO} , dotado da semântica $\langle +ac \rangle$;
- (d) p é a fonologia, $\langle \phi_i \rangle$, particular do núcleo; que permite, por exemplo, formar nomes em 'URA' e em 'EZA';
- (e) i e e são, respectivamente, os argumentos interno e externo do núcleo;
- (f) os três pontos, \dots , representam a lógica seletiva interna de h , aqui omitida.

A partir deste capítulo investigaremos como as estruturas sintáticas são derivadas, de modo sistemático, por meio de determinadas parametrizações do núcleo.

O verbalizador estativo, V_{BE} , por exemplo, é definido pela parametrização das quatro primeiras valências de h : (i) por se tratar de um verbo, o operador k recebe o valor v , próprio dos verbos; (ii) a semântica específica é estativa, $\langle +st \rangle$; (iii) a fonologia, neste caso, é *null* (nula), uma fonologia *default*, uma vez que a fonologia de V_{BE} é inteiramente gramatical, de-

terminada pelo operador v em combinação com o traço semântico $+st$; (iv) a função booleana, β , seleciona os argumentos interno e externo adequados à verbalização nominal, como veremos detalhadamente à continuação.

Portanto, o verbalizador estativo é definido pela parametrização de h :

$$V_{BE} = h \beta v \langle +st \rangle null.$$

Com as quatro primeiras valências preenchidas, h passa a possuir apenas as valências i e e abertas. Portanto,

$$V_{BE} = \lambda ie. (\lambda f. (f ((v \langle +st \rangle null) i)) e).$$

V_{BE} é um exemplo de *núcleo parametrizado*, ou, como passaremos a chamar, de *núcleo sintático*, por abuso de linguagem, uma vez que h é o núcleo sintático por excelência.

As várias parametrizações de h dão origem à sintaxe. Existem três tipos de núcleos sintáticos:

- (a) *Núcleos categorizadores*: R , o núcleo formador de raízes; N , o formador de nomes; e V , o formador de verbos.
- (b) *Núcleos modalizadores*: Asp , atribuidor de aspecto a um constituinte categorizado por V ou N ; $Tense$, responsável por situar um enunciado em relação ao enunciador;¹ Vox , caracterizador da voz de um constituinte verbalizado.
- (c) *Núcleos relacionadores*: Ve , formador de elementos vocabulares; $Comp$, associador de uma raiz a um constituinte nominal; associa uma raiz ao seu complemento; $Voice$, associador de um constituinte verbal a um constituinte nominal; associa um vP a um sujeito; $Cause$, associador de uma sentença a um constituinte nominal; associa uma sentença a um causador; $Prep$, várias adposições, associadoras de constituintes nominais.

Os núcleos listados acima compõem as estruturas sintáticas a serem analisadas nos próximos capítulos. A relação foi apresentada de modo extremamente sucinto, apenas para proporcionar uma visão de conjunto dos núcleos a serem abordados, uma vez que a caracterização desses elementos somente pode ser realizada em conjunto com as estruturas de que eles participam.

6.2 Relacionamento funcional entre os tipos de núcleos

A denominação *núcleo relacionador* é essencialmente inadequada, uma vez que todo núcleo relaciona dois termos, seus argumentos; entretanto, qualquer termo escolhido para designar esta classe de núcleos (*associador*, *juntor*, *vinculador*, etc.) padece do mesmo mal.

¹*Tense* será considerado como representante de uma classe de núcleos caracterizadores da relação existente entre um enunciado e o seu enunciador: *tempo* é uma das relações possíveis; há outras relações, como a *evidencialidade*, que caracteriza o enunciador como presente ao fato relatado ou se o ouviu de outro, etc..

Na base de qualquer derivação, como veremos, situa-se a junção de uma semântica, $\langle \sigma_i \rangle$, e uma fonologia, $\langle \varphi_i \rangle$, operação que se realiza por meio da aplicação $Ve \langle \sigma_i \rangle \langle \varphi_i \rangle$. Além de relacionar uma semântica e uma fonologia, o núcleo Ve também categoriza o constituinte resultante como um *elemento vocabular*. Ou seja, não há diferença fundamental entre um núcleo relacionador e um categorizador. A classificação de Ve como relacionador se deve ao modo como o núcleo entra em uma derivação, *à esquerda de seus dois argumentos*. Caso nossas derivações se iniciassem pela formação de raízes, o núcleo R passaria a ser classificado como relacionador, pois entraria em uma derivação aplicando-se a dois argumentos, *à esquerda* deles. Como será mostrado na continuação, os núcleos categorizadores e modalizadores entram em uma derivação *à direita* de um constituinte. Por exemplo, para o constituinte $\lambda f.(f(wa))b$ (o que quer que seja este constituinte) ser categorizado como raiz, *ele deve aplicar-se a R* , por meio da aplicação $(\lambda f.(f(wa))b)R$. Nesta aplicação, o núcleo R satura a valência f do constituinte, estabelecendo a relação $R(wa)b$, que é a categorização pretendida. Portanto, R , que entrou na derivação *à esquerda* em uma aplicação, acaba por situar-se *à direita* de dois valores, tornando-se o relacionador do conteúdo do constituinte: a diferença entre um núcleo relacionador e um categorizador ou modalizador é portanto contextual, e não “ontológica”.

Os dois modos como um núcleo pode entrar em uma derivação, *à esquerda* ou *à direita* de um termo, constituem a assimetria fundamental das derivações. Núcleos que entram na derivação *à esquerda* de um termo são operadores binários que acrescentam novos constituintes à derivação. Um exemplo é o núcleo *Comp* do próximo diagrama, que representa a derivação da sentença ‘A montanha é alta’, como futuramente veremos: ao aplicar-se à raiz rP , um termo previamente derivado, *Comp* leva à introdução de um nP , um constituinte nominal, como complemento da raiz. *Comp*, um núcleo aplicado *à direita* de rP , é responsável pelo acréscimo de material lexical, o nP , à derivação. Já o núcleo V_{BE} entra na derivação *à direita* de $\lambda f.(f(comp rP))nP$, como argumento deste constituinte formado pela ação de *Comp*. V_{BE} categoriza o constituinte como um verbo estativo. Isto é, V_{BE} acrescenta material funcional, o formador verbal, à derivação. Portanto, a assimetria de entrada dos núcleos em uma derivação, *à direita* ou *à esquerda* do constituinte situado no topo, é responsável pela agregação de material lexical e funcional à derivação.

do núcleo, sob o ponto de vista do fenômeno linguístico.

6.3.1 Evidência empírica

Nenhuma teoria linguística, atualmente, questiona o papel das raízes no processo gramatical de formação vocabular: nomes, verbos e adjetivos são reconhecidos como formados a partir de raízes. A principal questão atualmente consiste em decidir se a formação vocabular ocorre em um módulo próprio da gramática, separado do da formação sentencial, ou se vocábulos e sentenças são formados no mesmo sistema computacional. Segundo a minha pesquisa, o sistema único é praticamente inevitável e a fundamentação desta conclusão é tema de todo o texto, como veremos à continuação.

No entanto, há uma questão que precede a da unicidade do sistema computacional: até que ponto as raízes constituem o elemento mais básico da formação lexical?

Caso o sistema computacional seja realmente único, a indagação torna-se ainda mais crítica, pois questiona o próprio ponto de início das derivações sintáticas.

Acontece que os idiomas semíticos oferecem evidências muito fortes de que existe algo mais fundamental do que as raízes. Faremos uma análise do processo de derivação nominal e verbal no semita em geral (família de línguas semíticas, considerada como um todo) e em seguida investigaremos a existência de dois níveis de entidades participantes da derivação, utilizando dados do hebraico.

6.3.2 Evidência do semita

Em primeiro lugar, é preciso especificar que evidências devem-se buscar no semita para comprovação da proposta: (a) buscamos uma combinação de forma fonológica e significado que diretamente gere nomes e/ou verbos; a esta combinação, denomino *raiz*; (b) buscamos também uma combinação mais básica que dê origem à anterior; a esta formação mais básica, denomino *elemento vocabular*. É necessário caracterizar o alvo da busca, uma vez que não há uma terminologia geralmente aceita para tratar a formação morfo-sintática de palavras.

As línguas semíticas têm uma morfologia muito rica e o processo básico de formação de palavras é tradicionalmente definido como uma combinação de “raiz” e “*pattern*”.³

A raiz tipicamente consiste de três consoantes ordenadas, associadas a uma ideia geral ou significado básico. Um *pattern* consiste de uma sequência típica de vogais ou afixos que é aplicada às consoantes da raiz. Este processo forma as palavras efetivamente utilizadas na linguagem. Por exemplo, a raiz *ktb*, do árabe, tem o significado geral de “escrever”; o *pattern* CaCiC, aplicado a *ktb* deriva o nome ‘katib’, e o *pattern* CaCaCa, aplicado à mesma raiz, deriva o verbo ‘kataba’, “escrever”.

³Para as análises a seguir, utilizarei o termo “raiz” em sentido semi-formal, acompanhando o vocabulário habitual dos autores da área de morfo-sintaxe das línguas semíticas; não se deve confundir com o sentido do termo “raiz” formalizado neste trabalho.

O mesmo processo de *raiz e pattern* para a formação vocabular ocorre também nas línguas camíticas e há indícios bastante fortes de ter existido ainda como processo produtivo no sistema gramatical do proto-indo-europeu; de todo modo, o sistema verbal desta família linguística parece estar nele baseado. As línguas indo-europeias atuais exibem numerosos vestígios deste antigo método simbólico de flexão de raízes, principalmente nos verbos “fortes” e nos gramaticais; é o caso do inglês, com, por exemplo, ‘take’/‘took’, ‘sing’/‘sang’/‘sung’.

Com base no processo descrito, de modo bastante direto, a “raiz”, em geral trilítera, do semita é nominalizada ou verbalizada por um *pattern*. Ou seja, a raiz é o único elemento que serve de base ao processo de formação vocabular, não fornecendo, assim, a evidência que procuro de existência de um elemento mais básico do que a raiz, o *ve*. É necessário então mostrar que o termo resultante da aplicação de um *pattern* a uma raiz não é ainda estruturalmente uma palavra da linguagem. Caso seja possível fazê-lo, o processo de formação de palavras do semita proporciona a evidência procurada.

Para mostrar que, efetivamente, a “raiz” do semita corresponde ao meu *Ve* e o *pattern* corresponde ao meu *R* nos basearemos em dados do hebraico, de Maya Arad.

6.3.3 Evidência do hebraico

Arad (2002) “*argumenta por uma distinção entre a formação de palavras a partir de raízes e a formação de palavras a partir de outras palavras.*” Arad mostra que palavras derivadas de raízes apresentam uma ampla gama de interpretações, “*enquanto que palavras feitas de palavras existentes devem depender semântica e foneticamente das palavras das quais elas são derivadas.*”

Como é bem-conhecido, e é a base das análises de Arad, “*as três consoantes são impronunciáveis por si. O hebraico utiliza a morfologia de patterns para converter a raiz consonantal em uma palavra. A combinação de raízes e patterns serve a um duplo propósito: ela faz com que a raiz segmental se torne uma sequência pronunciável e transforma uma raiz (categorialmente neutra) em um nome, um verbo, ou um adjetivo.*” p. 3.

Depois de enunciar o processo de raiz e *pattern* de formação de palavras, Arad apresenta uma copiosa lista de exemplos ilustrativos de sua aplicação. Um exemplo particularmente claro é (p. 9):

- (24) raiz *qlt*
- a. CaCaC (verbo) ‘qalat’ (absorver, receber)
 - b. taCCiC (nome) ‘taqlit’ (um registro)

Portanto, a evidência é para uma clara formação do verbo ‘qalat’ a partir da raiz *qlt* e do *pattern* CaCaC, e do nome ‘taqlit’ a partir da mesma raiz em combinação com o *pattern* taCCiC.

Mas os exemplos podem ser investigados um passo mais adiante. Consideremos os exemplos da p. 17:

- (25) raiz *xšb* (atividade mental)
- a. hiCCiC (verbo) 'hexšiv' (considerar)
 - b. taCCiC (nome) 'taxšiv' (cálculo)
 - c. maCCeC (nome) 'maxšev' (computador)

Um *pattern*, usualmente (mas nem sempre), é composto de um prefixo e de um padrão de inserção de vogais. Os prefixos especificam ambiente nominais (nos exemplos, *ta* e *ma*) e verbais (*hi*, neste caso particular).

Os exemplos mostram a existência de duas formas: *xšiv* e *xšev*. A forma *xšiv* é verbalizada no ambiente *hi* e nominalizada no ambiente *ta*, a forma *xšev* é nominalizada no ambiente *ma*. O ponto importante, a ser observado, é que as formas *xšiv* e *xšev* são lexicalmente neutras, não dotadas de marcas fonológicas que as liguem a nomes ou verbos.

Assim, os dados do hebraico mostram evidências de um processo de categorização lexical que ocorre em duas etapas. Primeiramente, a inserção de vogais a uma sequência fonológica triconsonantal: o padrão CCiC aplicado à forma *xšb*, resultando *xšiv*; e o padrão CCeC, aplicado à mesma forma, produzindo *xšev*. Depois, a aplicação de prefixos às formas resultantes da etapa anterior, dando origem a nomes e verbos: a prefixação de *ta* e *ma* resulta nos nomes *taxšiv* e *maxšev*; e a prefixação de *hi* resulta no verbo *hexšiv*.

Segundo a gramática tradicional do hebraico, tem-se uma *raiz*, *xšb*, à qual se aplica um *pattern* (*binyan*). Nos termos do parágrafo anterior, a aplicação do *pattern* amalgama duas etapas.

No quadro geral da gramática do hebraico, a aplicação de *patterns* a raízes é um processo, em princípio, mais complexo do que o analisado acima. Pode-se verbalizar sem a ajuda de prefixos, como no caso do *pattern* CaCaC; e existem *patterns* ligados a formação da voz passiva, transitividade e causatividade verbais. Arad (2002, p. 7) mostra a raiz ʕmd dando origem a várias estruturas verbais, dependendo do *pattern* aplicado: CaCaC, 'ʕamad' (estar de pé); hiCCiC, 'heʕemid' (fazer levantar-se); huCCaC, 'huʕamad' (passiva da anterior). Essas palavras expressam estruturas verbais complexas por meio de flexões da raiz e prefixação.

Para as nossas análises, podemos deixar de lado os *patterns* relacionados à formação de vozes, transitividade, causação, por se tratar de fenômenos que ocorrem após a formação lexical. Processos gramaticais que em outras famílias linguísticas envolvem auxiliares e afixação à raiz no semita se realizam por flexão interna da raiz, combinada com afixos (que também podem ser *patterns*). Assim, a combinação de raiz e *pattern*, no semita, aplica-se a toda a estrutura da gramática. Vamos, assim, nos restringir ao processo de formação lexical.

Portanto, os dados do hebraico mostram evidências de três elementos participantes da

formação lexical:

- (a) Em primeiro lugar, há a associação de uma forma fonológica a um significado. A forma fonológica é composta por uma sequência de consoantes, em geral três (padrão CCC); o significado tem características de generalidade.
- (b) Sobre o padrão de consoantes, incide um padrão de inserção de vogais, produzindo como resultado uma forma fonológica adequada à pronúncia e um significado mais especializado.
- (c) À sequência fonológica anterior, aplica-se um afixo; resulta, então, uma palavra lexicalmente categorizada e um significado mais especializado.

Considero esses três elementos como evidências de um processo universal de formação lexical, que proponho como pertencente ao modelo computacional, de acordo com o mostrado a seguir.

- (a) O processo lexical básico é a simples união de uma forma fonológica a um significado geral. A este *par*, de forma fonológica e significado, denomino *elemento vocabular*; abreviadamente, *ve*. O processo que realiza a formação do par denoto por *Ve*, o qual, como todo processo gramatical, corresponde a uma parametrização especial do núcleo, *h*. *Ve* é o processo gramatical básico, o primeiro passo de uma derivação.⁴
- (b) Em seguida, o *ve* é submetido a um processo de especialização de significado e de alteração fonológica, sinalizadora do significado. Ao resultado da especialização do *ve*, denomino *raiz*.⁵ O processo que produz a raiz, a partir do *ve*, é denotado por *R*. O operador *R*, formador de raízes, é também um núcleo parametrizado.⁶
- (c) Finalmente, a raiz é lexicalmente categorizada, como *nome* ou como *verbo*. O processo de categorização é efetuado pelos operadores categorizadores *N* e *V*, respectivamente. Do mesmo modo que os operadores precedentes, *N* e *V* também são núcleos parametrizados. Os operadores *N* e *V* podem introduzir alterações sobre a fonologia da raiz e especializar o seu significado.

Por sua riquíssima morfologia, que inclui um processo produtivo de flexão interna de unidades lexicais, o semita permite que se observe de modo bastante explícito o processo universal de formação vocabular.

Entretanto, há ainda um passo faltante antes que se possa propor o processo recém-analisado como uma base universal para a formação lexical: um *ve* e uma raiz são destituídos de categorização nominal ou verbal, nada pode haver neles que possibilite ao sistema computa-

⁴E toda derivação é sintática, segundo a proposta.

⁵O processo de generalização/especialização, presente em um par *ve/raiz*, pode ser revelador de uma capacidade primitiva humana, de natureza cognitiva.

⁶Como veremos posteriormente, a fonologia do elemento vocabular não é modificada; a fonologia da raiz é uma “imagem” da fonologia do *ve*. A obtenção da fonologia da raiz é semelhante a um resultado de operação matemática: quando se diz “o dobro de 2 é 4”, obtém-se o resultado 4, mas o argumento 2 permanece inalterado.

cional tratá-los diretamente como nomes ou verbos.

No entanto, voltando aos dados do hebraico: o *pattern* CaCaC, que, por hipótese, é um formador de raízes, caracteriza diretamente a raiz trilítera qlt como verbo — ‘qalat’ (absorver, receber). Caso este exemplo seja interpretado *prima facie*, o *pattern* é portador de uma categorização verbal e a proposta do processo universal de formação lexical cai por terra.

O *pattern* CaCaC é um exemplo do fenômeno translinguístico das raízes “verbais” e “nominais”. Em português, a raiz de fonologia abr, da qual se deriva o verbo ‘abrir’ é uma raiz dita verbal, e a raiz fac, originadora do nome ‘faca’ é uma raiz “nominal”. Uma tal classificação ocasiona dificuldades com relação às raízes de ‘baile’ e ‘bailar’, e de ‘dança’ e ‘dançar’: são nominais ou verbais? Por convenção, se o sentido básico da raiz for o de *atividade*, a raiz é verbal; caso contrário é nominal. O critério é um tanto arbitrário, mas não deixa de ter fundamento.

No caso do *pattern* CaCaC, verifica-se que ele não é exclusivo da formação de verbos: aplica-se à raiz trilítera xzq para produzir ‘xazaq’ (forte), um adjetivo. Portanto, CaCaC, por si, não é um categorizador (verbalizador ou formador de adjetivos).

Nos termos da minha proposta, a raiz trilítera xzq é um elemento vocabular, e, como tal, é isenta de categorização lexical. Uma vez que o *pattern* CaCaC também não porta uma tal categorização, como acabamos de comprovar, o caráter adjetival de ‘xazaq’ só pode provir da semântica do *ve* xzq, uma semântica claramente estativa; isto é, dentro do significado de xzq está o traço semântico $\langle +st \rangle$. Ocorre que o traço $\langle +st \rangle$ é tão-somente um traço semântico, não é um categorizador lexical. Deste modo, nem o *pattern*, nem o *ve* são portadores de qualquer característica computacional capaz de gerar um termo categorizado como adjetivo ou verbo.

Portanto, e este é um fato fundamental, a categorização só pode ser resultado de um categorizador atuante *após* a aplicação do *pattern* ao elemento vocabular. Ou seja, há um processo de categorização lexical que se aplica inclusive ao *pattern* CaCaC, ocorre entretanto que o processo apresenta um acréscimo fonológico nulo, um fenômeno linguístico comum.

Mantém-se, assim, a proposta de que a formação lexical compõe-se de três operações, três núcleos: *Ve*, *R*, *V* ou *N*. Existem também os adjetivos, mas a este tema voltaremos posteriormente.

A divisão acima corresponde a uma propriedade fundamental dos núcleos, segundo a minha proposta: *cada núcleo introduz apenas um efeito sobre o termo a que se aplica*. O *Ve* reúne uma forma fonológica e um significado, e somente isto. O núcleo *R* prepara a junção assim formada, fonológica e semanticamente, para participar de uma derivação. Uma raiz deve, necessariamente, apresentar estrutura fonológica e semântica adequadas à formação de expressões linguísticas, segundo as exigências especiais de uma determinada língua; a raiz exige, assim, um processo particular de formação. Esta preparação é uma segunda operação, que se segue à simples junção de uma forma fonológica e um significado. Uma vez formada, a raiz deve ser qualificada como verbo (operador *V*) ou como nome (operador *N*), as duas categorias lexicais que compõem uma enunciação. E esta categorização é um terceiro efeito.

O modelo computacional não exige que uma derivação se inicie a partir de um ele-

mento vocabular; as operações gramaticais poderiam igualmente bem começar tomando por base uma raiz. De todo modo, considero o núcleo *Ve* como um primitivo linguístico: onde quer que haja um efeito especial, ali haverá um núcleo específico que o realiza.

6.4 Resumo provisório

Temos, então, quatro núcleos de formação lexical: *Ve*, *R*, *N* e *V*. Os quatro núcleos, como veremos detalhadamente à continuação, geram constituintes de estrutura

$$\lambda f.(f(g\langle\sigma_i\rangle))\langle\varphi_i\rangle,$$

em que *g* é um operador simples (no caso do *Ve*) ou composto (nos demais casos). Ou seja, os constituintes lexicais possuem um componente semântico, $\langle\sigma_i\rangle$, e um fonológico, $\langle\varphi_i\rangle$, sob a ação de um operador, *g*.

Os núcleos responsáveis pela inclusão de traços semânticos e fonológicos, em seu nível mais básico, são *Ve* e *R*. Partirei da hipótese (hipótese compartilhada com a MD) de que a semântica mínima necessariamente codificada por um elemento vocabular é a de *estatividade* ou de *atividade*, mas não ambas simultaneamente. Isto é, um elemento vocabular é necessariamente estativo ou ativo; sua sequência mínima de traços semânticos é, assim, $\langle+st\rangle$ ou $\langle+ac\rangle$.⁷ O núcleo *R* acrescenta outros traços semânticos àqueles já incluídos por *Ve*, tipicamente, *in*, *ec* e *ic*, qualificados como *necessários*+, *proibidos*- ou *opcionais*.

Consideremos, por exemplo, o verbo *cortar*. A análise da estatividade ou atividade um item básico, como um *ve* ou uma raiz, requer a análise dos ambientes sintáticos em que o item pode figurar, tema dos próximos capítulos. Para efeitos imediatos, considerarei que *cortar* denota uma atividade, com base principalmente no fato de a raiz *cort* exibir estatividade derivada (*cortado/a*), diferentemente do par *azedar/lazedo*. Ou seja, a estatividade não é uma propriedade primitiva da raiz *cort*, mas de um ambiente verbal em que ela está inserida. Resta portanto para esta raiz o traço semântico $+ac$, por sinal compatível com uma semântica ingênua do verbo *cortar*, costumeiramente classificado como verbo ‘de ação’.

Com base nos pressupostos adotados, pode-se razoavelmente supor que tal verbo tem origem em um *ve* de semântica lexical $\langle+ac\rangle$ e fonologia *cort*. Com esta composição semântica e fonológica, o *ve* é o constituinte

$$\lambda f.(f(ve\langle+ac\rangle))\text{cort.}$$

A raiz derivada deste *ve* nada altera à semântica ou à fonologia do *ve*, mas deve codificar os fatos de: a atividade incidir sobre algo (traço semântico $+in$); ter uma causação necessariamente externa ao evento denotado (traço $+ec$); e, em consequência do traço anterior, ser isenta de

⁷Ver p. 47 para a conceituação dos traços semânticos.

causação interna ao evento (traço *-ic*). Uma tal raiz é codificada pelo constituinte

$$\lambda f.(f(r(ve\langle +ac, +in, +ec, -ic\rangle)))\text{ cort.}$$

Os dois exemplos acima são meramente ilustrativos para exemplificar o relacionamento entre os operadores (*r* e *ve*) e os componentes semânticos e fonológicos; a derivação dos constituintes mostrados é assunto da próxima seção.

Além dos componentes semântico e fonológico, cujo conteúdo é estabelecido pelo sistema gramatical, um constituinte lexical tem também um *significado enciclopédico*, que reúne as ideias associadas ao constituinte, de modo assistemático, difuso e não codificados pela gramática. Por exemplo, o significado enciclopédico de *gato* inclui os conceitos de felino, quadrúpede, arredo, habituado a subir em telhados, etc., nenhum dos quais é codificado pelo sistema gramatical da língua portuguesa.⁸ Qualquer constituinte pode ter um significado enciclopédico, não apenas os constituintes lexicais: a sentença ‘Maria vai com as outras’ possui o significado enciclopédico de “pessoa que não tem opiniões próprias”, entre outros. Por razões programáticas, o significado enciclopédico será deixado à margem da pesquisa; ele será usado apenas nos casos específicos em que a estrutura dos constituintes não for suficiente para determinar os resultados pretendidos, ou seja, para estabelecer os limites do modelo proposto.

O componente semântico e, em parte, o fonológico, são responsáveis pelo comportamento gramatical particular do constituinte; diferenciam, por exemplo, *pulo*, *pular* e *cair*. Grande parte da pesquisa tem em vista estabelecer a ligação entre os componentes semânticos e o comportamento gramatical dos constituintes. Pretendo mostrar como os traços semânticos *ac*, *st*, *in*, *ec* e *ic*, em combinação com núcleos, engendram a estrutura argumental nominal e verbal; como características de *atividade/estatividade*, *incidência* e *causalidade* dão origem a *sujeitos*, *objetos diretos* e *indiretos* (caracterizados puramente pela estrutura dos constituintes, como veremos) e às variações e alternâncias entre eles.

Não se trata apenas de produzir formas a serem pronunciadas, como ‘pulo’ ou ‘pulou’, mas de gerar famílias de estruturas como:

- (26) O repolho azedou
- (27) Maria azedou o repolho
- (28) O macaquinho engavetou-se
- (29) João pulou
- (30) João pulou o muro

⁸O significado enciclopédico refere-se à semântica não codificada pelo sistema linguístico, e não às categorias com que lidam lógicos ou semanticistas, ainda que sobre base científica. Para a língua, *ser quadrúpede ou felino* não é “uma propriedade necessária dos gatos”, mas um fato empírico que determinados falantes podem mesmo desconhecer. Para uma apresentação informada do assunto pode-se consultar Eco (1991a, pp. 63–135), em que se conceituam os registros enciclopédicos como “*conteúdo convencional da expressão*”, diferentemente de “*necessidade lógica*”.

(31) O pulo do João

Teorias que postulam grades temáticas para itens lexicais têm grandes dificuldades para lidar com as alternâncias e as variações argumentais exemplificadas acima. No modelo computacional que proponho não há como especificar grades temáticas de modo natural; a derivação de estruturas a partir de poucas características semânticas muito gerais e a interpretação estrutural dos constituintes derivados decorrem do modelo, embora o modelo imponha rigorosas restrições e a obrigação de trabalhar com recursos deveras escassos.

6.5 Derivação lexical básica

Passemos à derivação das estruturas sintáticas, começando pela derivação do nome *beleza*, percorrendo todos os passos derivacionais, a partir do elemento vocabular, passando pela raiz.

O primeiro passo da derivação consiste da determinação das características semânticas e fonológicas do elemento vocabular e da raiz.

O elemento vocabular é claramente estativo e possui a fonologia *bel*;⁹ seus traços semânticos e fonológicos são, assim, determinados sem dificuldade: *+st* e *bel*.

Determinar as características da raiz, no entanto, demanda alguma análise, que leve em conta determinados contextos em que a raiz pode figurar. Além de *beleza*, a raiz também participa da formação de *bela* e *belo*, duas formas de um adjetivo. Não há indícios de que a raiz participe diretamente da formação de verbos. Evidentemente, há o verbo *embeleazar*, mas aí a raiz aparece em combinação com outros elementos, o que mostra que formar verbos não é a característica básica da raiz. Uma vez que *beleza*, um nome, também exhibe a interferência de outros fenômenos gramaticais, pode-se concluir que a semântica básica da raiz é a de “qualidade”, fato compatível com o traço *+st*. Portanto: (a) trata-se de uma condição de algo; é portanto uma raiz que requer incidência, ou seja, o traço *+in*; (b) o fato de alguém, ou algo, ser belo, em princípio, é uma condição própria, independente de ações externas; tem-se, assim, os traços *–ec* e *+ic*.¹⁰ A raiz, então, é dotada da fonologia *bel*, adquirida do elemento vocabular, e dos traços semânticos $\langle +st, +in, -ec, +ic \rangle$. *Aparentemente*, não existe dificuldade para a formação do nome, *beleza*: em uma primeira aproximação (a ser revista alguns parágrafos à frente), há o acréscimo de um sufixo (*eza*) à raiz; e os traços semânticos não parecem ir além dos da raiz.

Estamos, assim, equipados com os elementos semânticos e fonológicos necessários às derivações que levem à formação do nome pretendido.

⁹Não confundir o *caráter estativo* do elemento vocabular com “ser um estado”, o que um *ve* não pode ser.

¹⁰A determinação dos traços semânticos da raiz, na realidade, depende de muitos outros fatores. É necessário, principalmente, levar em conta a compatibilidade com diversas estruturas verbais (‘Maria é bela’), como a possibilidade das vozes passiva e média, alternâncias de transitividade, construções causativas, etc.. No momento não há como invocar o arsenal disponível, o que irá sendo desenvolvido, progressivamente, ao longo do trabalho. O parágrafo apenas procura dar uma mostra da metodologia de análise a ser utilizada.

6.5.1 Derivação do elemento vocabular

Qualquer núcleo é definido por uma parametrização específica de h , o núcleo universal. Devemos lembrar que h é definido por

$$h = \lambda b k s p i e. (\lambda f. (f ((k s p) i)) e).^{11}$$

O elemento vocabular apenas reúne um feixe de traços semânticos e uma sequência de traços fonológicos, sem lhes acrescentar qualquer outra semântica ou fonologia. Assim, o núcleo Ve é definido pela seguinte parametrização de h :

$$Ve = h \beta ve null null,$$

em que: (a) ve é o operador característico do núcleo; (b) os dois valores *null* indicam os acréscimos semântico e fonológico do núcleo; isto é, não há acréscimos semânticos ou fonológicos; (c) β é a função booleana que valida os argumentos interno e externo do núcleo, i e e , ainda a serem introduzidos.

Com as quatro primeiras valências de h preenchidas, Ve , então, possui somente as posições argumentais i e e em aberto. Portanto,

$$Ve = \lambda i e. (\lambda f. (f ((ve null null) i)) e).$$

Esta é a forma mais adequada do núcleo Ve para uso em derivações; o núcleo é especificado como um constituinte em que os argumentos interno e externo se encontram em aberto.

Utilizando-se a especificação acima, a derivação do elemento vocabular de traço semântico $+st$ e fonologia bel é imediata:

$$Ve \langle +st \rangle bel = (\lambda i e. (\lambda f. (f ((ve null null) i)) e)) \langle +st \rangle bel.$$

Substituindo-se i e e pelos seus valores, resulta:

$$Ve \langle +st \rangle bel = \lambda f. (f ((ve null null) \langle +st \rangle)) bel.$$

Temos, acima, a estrutura do item vocabular de fonologia bel e traço semântico $+st$, expressa de dois modos equivalentes: à esquerda, em formato de aplicação funcional; à direita, em formato de constituinte.

Neste ponto, será usada uma definição para facilitar a visualização da expressão. No caso do núcleo Ve , o operador ve , introduzido pelo núcleo, somente pode ter argumentos nulos;

¹¹Para simplificar a análise, a definição contempla apenas o processamento bem-sucedido de h , deixando de lado a condição de erro. O leitor, no entanto, deve ter em mente que uma aplicação de h a valores inaceitáveis produz o término anormal da derivação.

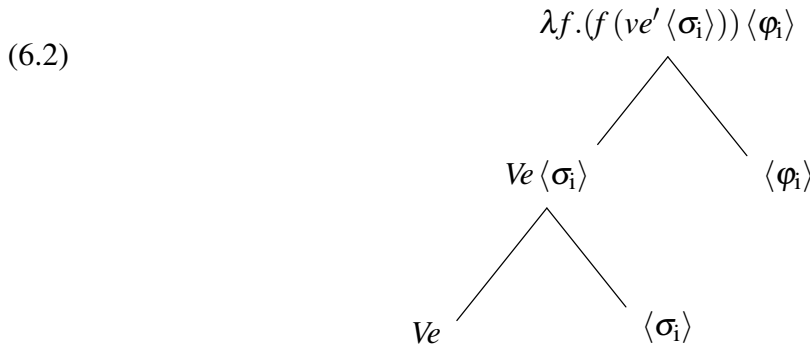
por tal motivo, a aplicação $ve\ null\ null$ será definida pelo operador de mais alto nível ve' ; assim,

$$ve' = ve\ null\ null.$$

Com esta simplificação, o elemento vocabular é expresso por

$$Ve\ \langle +st \rangle\ bel = \lambda f.(f(ve'\ \langle +st \rangle))\ bel.$$

O diagrama a seguir representa a derivação de um item vocabular no caso geral, em que os traços semânticos e a fonologia recebem uma definição qualquer.



Argumentos exteriores do elemento vocabular

Consideremos o núcleo parametrizado Ve expresso sob o formato

$$Ve = \lambda ie.(\lambda f.(f((ve\ null\ null)\ i))\ e).$$

Segundo as análises da p. 55 e seguintes e da p. 60, o operador ve introduz o paradigma gramatical w_{ve} , responsável pela formação dos elementos vocabulares; utilizando-se explicitamente o paradigma gramatical, Ve é expresso por

$$Ve = \lambda ie.\lambda f.(f(\lambda c.(c(w_{ve}\ null\ null))\ i))\ e,$$

um núcleo que apresenta os parâmetros i e e em aberto, a serem preenchidos, respectivamente, por um feixe de traços semânticos e um feixe de traços fonológicos, de acordo com o conceito de elemento vocabular. Aplicando Ve a uma semântica e uma fonologia quaisquer, resulta o constituinte maximal

$$\lambda f.(f(\lambda c.(c(w_{ve}\ null\ null))\ \langle \sigma_i \rangle))\ \langle \phi_i \rangle.$$

A questão é determinar, neste constituinte maximal, qual é o tipo dos dois argumentos do contexto c , a saber, qual é o tipo dos argumentos $(w_{ve}\ null\ null)\ \langle \sigma_i \rangle$ e $\langle \phi_i \rangle$. A determinação desses tipos é importante, pois condiciona a especificação da função de seleção dos núcleos que virão

a processar um elemento vocabular, em especial a do núcleo R formador de raízes.

Considerando-se que o argumento fonológico $\langle \phi_i \rangle$ não é afetado pelo paradigma w_{ve} , não se produzindo assim uma alteração em seu tipo, a questão se reduz a determinar qual é o tipo da computação $(w_{ve} \text{ null null}) \langle \sigma_i \rangle$.

De acordo com as análises da p. 61, esta computação produz como resultado um feixe de traços semânticos, uma vez que o argumento interno do operador composto $w_{ve} \text{ null null}$ é $\langle \sigma_i \rangle$, um feixe de traços semânticos.¹²

Os dois argumentos do contexto c são portanto feixes de traços semânticos e fonológicos, nesta ordem. Como c é indeterminado enquanto operador interno de Ve , o único conteúdo de informação que um elemento vocabular tem a oferecer para as funções de seleção de núcleos situados acima dele na derivação são um feixe de traços semânticos e um feixe de traços fonológicos.

6.5.2 Núcleos e traços semânticos

A questão efetivamente é mais ampla do que a colocada pela discussão dos parágrafos anteriores: na aplicação $ve' \langle \sigma_i \rangle$, o modelo computacional determina que o operador ve' ao aplicar-se aos traços $\langle \sigma_i \rangle$ *pode* alterar os traços semânticos processados, desde que o resultado continue a ser um feixe de traços semânticos.

Considerando a questão mais concretamente: na derivação do elemento vocabular de fonologia *bel* anteriormente exemplificado, a aplicação $ve' \langle +st \rangle$ não está proibida de produzir um resultado diferente de $\langle +st \rangle$, desde que o resultado seja ainda um feixe de traços semânticos. Tal fato significa que a aplicação funcional $ve' \langle \sigma_i \rangle$ pode ser tratada como geradora do feixe de traços semânticos $\langle t_{ve}, +st \rangle$, em que t_{ve} é um traço semântico denotador de elementos vocabulares. Neste caso, a simples combinação da sequência de traços semânticos $\langle t_{ve}, +st \rangle$ com uma fonologia qualquer já se caracteriza como um item vocabular, sem que para tanto seja necessária a presença de qualquer núcleo.

Assim, a técnica de se redefinir um núcleo em termos de operadores de mais baixo nível é substituída pelo recurso alternativo de se fazer um tal núcleo gerar um traço semântico especial, incluído convenientemente em um feixe de traços. Dito de outro modo: é possível trocar um modelo computacional em que as derivações deixam explícitos os operadores, como ve' , preservando a estrutura funcional dos núcleos, por outro modelo em que os operadores são substituídos por traços semânticos, como t_{ve} (na realidade operadores semânticos). Em outras palavras: as derivações podem ser representadas sintaticamente, por meio de composição de operadores, ou semanticamente, por meio de listas de traços; no primeiro caso, tem-se a derivação representada em um formato derivacional; no segundo caso, a derivação exibe uma aparência configuracional. Os dois estilos de representação são equivalentes no sentido de

¹²Tem-se um operador composto $w_k sp$ aplicado a um termo t , produzindo como resultado um termo t' , do mesmo tipo de t (ver p. 61).

serem capazes de gerar as mesmas estruturas.

Neste trabalho, apenas empregaremos o modelo estritamente sintático, preservando os núcleos ao longo das derivações, uma vez que assim se opera com menor quantidade de elementos (apenas os núcleos); o modelo semanticamente orientado, embora eventualmente proporcione algumas facilidades de implementação, demanda mais ampla investigação, não podendo entretanto ser descartado.

É possível levar a questão um pouco mais adiante: mesmo os traços semânticos considerados primitivos — como *st*, *ac*, *in*, *ec* e *ic* — podem ser reinterpretados em termos de núcleos, recebendo assim um tratamento sintático. Sequências como $\langle +st, +in, ic \rangle$ deixariam de ser utilizadas pela sintaxe como objetos dados, previamente formados em um sistema computacional próprio; a reinterpretação implica que uma sequência primitiva de traços seja substituída por uma composição de núcleos, cada núcleo exercendo a função de um traço semântico.

Como considero que a semântica universal (aquela minimamente composta pelos traços *st*, *ac*, *in*, *ec* e *ic*, além de possíveis outros) realmente constitui um sistema computacional separado do da sintaxe sentencial-lexical, os traços semânticos universais e suas sequências continuariam a ser tratados como termos primitivos pelo modelo computacional deste trabalho. Não há assim motivos para a implementação dos traços semânticos como núcleos, aplicando-lhes um tratamento sintático: incrementar-se-ia a complexidade do modelo computacional sob investigação, sem que houvesse ganhos suplementares que o justificassem.

6.5.3 Derivação da raiz

A definição do núcleo *R*, formador de raízes, segue o processo canônico de parametrização do núcleo universal *h*. Entretanto, com uma importante diferença em comparação com a parametrização de *Ve*. Para o formador de elementos vocabulares, as posições argumentais *s* e *p* são sempre nulas; já para *R*, as mesmas posições são praticamente imprevisíveis, sobretudo quanto ao acréscimo de traços semânticos. Assim, no caso geral, a parametrização que especifica *R* só ocorre para as posições argumentais *b* e *k*. Tem-se, então;

$$R = h \beta r,$$

em que: (a) β seleciona um feixe de traços semânticos para *i*, o argumento interno, e um feixe de traços fonológicos para *e*, o argumento externo;¹³ (b) *r* é o operador característico do núcleo. Com esta parametrização, temos

$$R = \lambda s p i e. (\lambda f. (f ((r s p) i)) e),$$

em que a semântica, *s*, e a fonologia, *p*, acrescentadas pelo núcleo, permanecem em aberto.

¹³*R* tem em vista aplicar-se ao conteúdo de informação de um elemento vocabular (ver p. 79), fato que justifica o critério de seleção de β .

No caso específico da raiz que queremos representar, a ser apresentada algumas linhas abaixo, o acréscimo semântico da raiz consiste da sequência de traços $\langle +in, -ec, +ic \rangle$, e o acréscimo fonológico é nulo. O formador de raízes que temos em vista pode então ser obtido por meio de uma segunda parametrização:

$$R\langle +in, -ec, +ic \rangle null = \lambda ie.(\lambda f.(f((r\langle +in, -ec, +ic \rangle null) i)) e).$$

O operador, acima parametrizado, é inteiramente funcional, porém é pouco prático para uso em derivações, devido à sua complexidade visual. É possível reduzi-lo a um formato mais manejável, do seguinte modo:

- (a) A sequência de traços semânticos $\langle +in, -ec, +ic \rangle$ é característica das construções inacusativas, em que: (i) um argumento introduzido para satisfazer necessidades semânticas da raiz (traço $+in$) exerce uma causalidade interna necessária (traço $+ic$) relativa ao evento denotado pela raiz; e (ii) não há causalidade externa relacionada ao mesmo evento (traço $-ec$).
- (b) A raiz não introduz complementações aos traços fonológicos pertencentes ao elemento vocabular; ou seja, a posição argumental p é saturada com o valor *null*.

Deste modo, pode-se “empacotar” a aplicação $R\langle +in, -ec, +ic \rangle null$ por meio da definição

$$R_{INAC} = R\langle +in, -ec, +ic \rangle null.$$

O núcleo parametrizado R_{INAC} é então o formador de raízes inacusativas.

Analogamente, o operador $r\langle +in, -ec, +ic \rangle null$ pode ser definido como

$$r_{INAC} = r\langle +in, -ec, +ic \rangle null.$$

Utilizando-se essas definições, o formador de raízes é então expresso por

$$R_{INAC} = \lambda ie.(\lambda f.(f(r_{INAC} i)) e).$$

O processo de tornar mais legíveis as expressões, por meio de definições, corresponde ao ato de lexicalização de constituintes: uma definição equivale à formação de um item lexical que passa a ser utilizado com uma estrutura e um significado previamente definidos.

Estamos agora preparados para categorizar o elemento vocabular

$$Ve\langle +st \rangle bel = \lambda f.(f(ve'\langle +st \rangle)) bel$$

como raiz.

A categorização de um constituinte envolve a execução de operações sobre os seus componentes internos, semântico e fonológico. Contudo, todo constituinte possui a estrutura $\lambda f.(f(ga))b$, em que a variável f oferece a possibilidade de que uma operação se “infiltre”

para operar sobre o conteúdo interno, (ga) e b . Exatamente esta variável será utilizada pelo núcleo R_{INAC} para categorizar o elemento vocabular: R_{INAC} será “absorvido” pelo elemento vocabular por meio da aplicação funcional

$$(\lambda f.(f(ve' \langle +st \rangle))) bel) R_{\text{INAC}}.$$

Ao substituir a variável f , R_{INAC} ocupará a posição necessária a operar sobre os dois argumentos, $ve' \langle +st \rangle$ e bel , produzindo-se a operação categorizadora

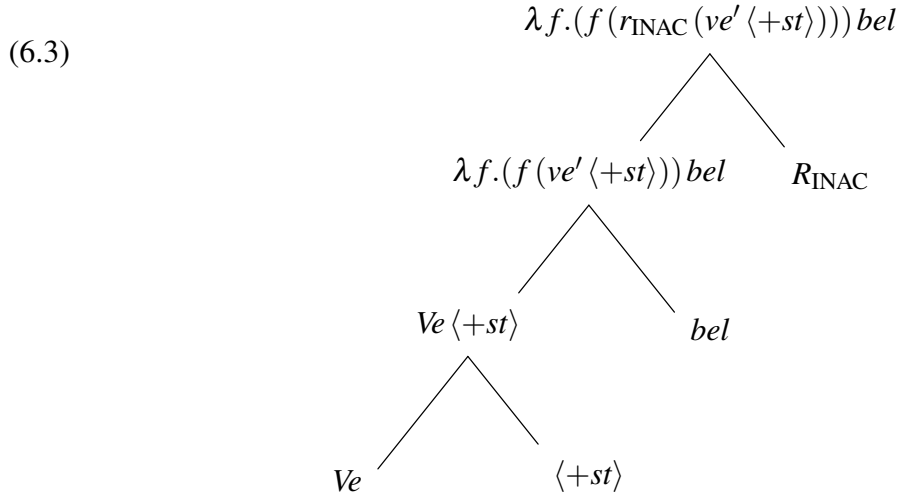
$$R_{\text{INAC}}(ve' \langle +st \rangle) bel.$$

A computação desta aplicação funcional produz a esperada categorização do elemento vocabular como raiz; a raiz recém-categorizada é o constituinte

$$\lambda f.(f(R_{\text{INAC}}(ve' \langle +st \rangle))) bel,$$

uma raiz inacusativa.

A derivação da raiz é mostrada no diagrama a seguir.



O constituinte categorizado como raiz, situado no topo do diagrama, codifica toda a sua história derivacional em sua própria estrutura, de modo explícito: lá estão os parâmetros lexicais $(+st)$ e bel) e os operadores segundo a ordem de inclusão na derivação (primeiramente ve' , depois r_{INAC}); lá está também a semântica particular dos operadores, como a exibida por r_{INAC} . A complexidade de um constituinte corresponde à complexidade da sua derivação.

6.5.4 Derivação do nome

Do mesmo modo que para o núcleo R , o nominalizador genérico N é definido pela parametrização adequada das variáveis b e k de h , como mostrado abaixo:

$$N = h\beta n,$$

uma expressão equivalente a

$$N = \lambda spie.(\lambda f.(f((nsp)i))e).$$

Analisamos anteriormente a geração dos nomes em *ura* (seção 5.5, p. 48), como exemplo de parametrização do núcleo universal. A especificação do núcleo N_{URA} então empreendida teve finalidade meramente ilustrativa, deixando à margem alguns ingredientes da derivação. Abordaremos agora a derivação nominal em suas três etapas (derivação do elemento vocabular, da raiz e do nome). Iniciaremos pela derivação do nome a ser pronunciado como ‘beleza’, para exemplificar a formação de uma classe de nomes, obtida por uma parametrização específica do núcleo nominalizador N ; em seguida faremos a derivação de um nome no caso mais geral.

A derivação do nome pronunciado como ‘beleza’ resulta da sufixação da sequência fonológica *eza* à raiz de fonologia *bel*. De acordo com o modelo computacional, o nome é gerado pelo núcleo N , que, no caso, introduz uma contribuição fonológica, *ez*, e uma contribuição semântica, $\langle sing, fem \rangle$, ao nominalizar uma raiz de fonologia *bel*. Computacionalmente, a formação de nomes “em *eza*” é semelhante à dos nomes “em *ura*”, previamente analisada.

O nominalizador que acrescenta o sufixo *eza* apresenta uma atuação de ampla generalidade, formando substantivos abstratos a partir de determinadas raízes estativas; ou seja, trata-se de um nominalizador com propriedades a serem capturadas por um núcleo especial. A especificação do nominalizador é feita por uma parametrização de N , em que se atribuem valores adequados às variáveis s e p , como a seguir:

$$N \langle sing, fem \rangle ez = \lambda ie.(\lambda f.(f((n \langle sing, fem \rangle ez)i))e).$$

A parametrização $N \langle sing, fem \rangle ez$ requer comentários.

Primeiramente, deve-se observar que a formação do nome resulta do operador n : o paradigma gramatical interno ao operador n interpreta a sequência fonológica *ez* como ligada à formação de nomes abstratos e determina que a fonologia do núcleo deve ser expressa por meio de sufixação; a partir dos traços semânticos $\langle sing, fem \rangle$ é gerada então a sequência fonológica *eza* a ser sufixada à fonologia da raiz.

Tal processo deixa ainda uma importante questão a ser analisada: se o paradigma gramatical interno a n interpreta a fonologia *ez* como indicadora de nomes abstratos e decide pela

sufixação da fonologia resultante, por que então não atribui uma classificação semântica *default*, *sing* e *fem*, ao sufixo, gerando a sequência fonológica *eza*? Não seriam assim necessários os traços semânticos $\langle sing, fem \rangle$ parametrizadores de N ; a expressão fonológica do núcleo seria inteiramente gramatical, como o é a semântica.

Esta é uma possibilidade a ser seriamente considerada, mas há alguns fenômenos que a tornam problemática.

Em primeiro lugar, é perfeitamente possível a formação gramatical do nome 'belezas', evidenciando que a expressão de número não pertence a um *default* gramatical, o qual ficaria, por consequência, restrito apenas ao traço semântico *fem*. Há entretanto nominalizadores para os quais este traço não é determinado pelo paradigma gramatical. Pode-se especificar, por exemplo, um núcleo formador de nomes indicadores de ocupação em *eiro/a* ('engenheiro'/'engenheira'), em que o traço *fem* não é gramaticalmente determinado.

Deste modo, aquilo que se denomina a "expressão gramatical" de um núcleo (como a expressão do traço $+fem$ de N_{URA} e de N_{EZA} a ser definido) deriva do processo de lexicalização (parametrizações de núcleos registradas no léxico, sendo portanto aprendidas) em combinação com o sistema computacional universal (responsável pela interpretação e composição de traços).

Consideradas as razões acima, as contribuições semânticas e fonológicas de núcleos que contêm elementos não-universais (como o traço semântico *fem* e a fonologia *eza*) serão sempre especificadas por meio dos parâmetros s e p de h , sendo sempre consideradas externas aos operadores gramaticais, como n e v . Deve-se ter em conta que os parâmetros interiores, como s e p , destinam-se justamente a organizar a lógica interna de um núcleo, naquilo que ela tem de variável.

Podemos então prosseguir a derivação a partir do núcleo parametrizado $N \langle sing, fem \rangle ez$, que codifica um comportamento gramatical específico que pode ser gramaticalizado por meio da definição de núcleo

$$N_{EZA} = N \langle sing, fem \rangle ez,$$

e da correspondente definição de operador

$$n_{EZA} = n \langle sing, fem \rangle ez.$$

Ou seja,

$$N_{EZA} = \lambda ie. (\lambda f. (f (n_{EZA} i)) e).$$

Deste modo, o nome expresso fonologicamente como 'beleza' resulta da nominalização da raiz de estrutura $\lambda f. (f (r_{INAC} (ve' \langle +st \rangle))) bel$ por meio do núcleo N_{EZA} . Isto é,

$$(\lambda f. (f (r_{INAC} (ve' \langle +st \rangle))) bel) N_{EZA},$$

cujo resultado é o constituinte nP

$$\lambda f.(f(n_{EZA}(r_{INAC}(ve' \langle +st \rangle)))) bel.$$

Comumente, um nP como *beleza* é considerado como um nome derivado de um adjetivo; segundo o modelo proposto, trata-se da derivação de um nome a partir de uma raiz estativa. O *status* dos adjetivos será futuramente analisado. A derivação do nP constituinte encontra-se descrita no diagrama a seguir.

(6.4)

$$\begin{array}{c} \lambda f.(f(n_{EZA}(r_{INAC}(ve \langle +st \rangle)))) bel \\ \swarrow \quad \searrow \\ \lambda f.(f(r_{INAC}(ve \langle +st \rangle))) bel \quad N_{EZA} \end{array}$$

A estrutura do constituinte situado no topo do diagrama mostra, de forma explícita, a história derivacional do nP : (a) uma semântica, $\langle +st \rangle$, e uma fonologia, *bel*, foram sucessivamente afetadas por três processos; (b) a formação de um elemento vocabular, por ve' ; a categorização como raiz inacusativa, por r_{INAC} ; a categorização como nome de um determinado tipo, por n_{EZA} .

O contraste entre os núcleos categorizadores, R_{INAC} e N_{EZA} , e o núcleo criador de associação, Ve , é destacado na versão completa da derivação, exibida no diagrama abaixo: os primeiros são incluídos na derivação à sua direita, e o segundo, à esquerda.

(6.5)

$$\begin{array}{c} \lambda f.(f(n_{EZA}(r_{INAC}(ve \langle +st \rangle)))) bel \\ \swarrow \quad \searrow \\ \lambda f.(f(r_{INAC}(ve \langle +st \rangle))) bel \quad N_{EZA} \\ \swarrow \quad \searrow \\ \lambda f.(f(ve \langle +st \rangle)) bel \quad R_{INAC} \\ \swarrow \quad \searrow \\ Ve \langle +st \rangle \quad bel \\ \swarrow \quad \searrow \\ Ve \quad \langle +st \rangle \end{array}$$

6.5.5 Derivação genérica de nome

Na seção anterior, analisamos a formação de uma determinada classe de nomes; passaremos, agora, ao caso geral, em que não há o acréscimo de um significado específico à semântica da raiz.

Consideraremos nomes como *banca* ou *banco*, em que uma raiz é diretamente nominalizada. Mediante uma análise sumária, pode-se concluir pela existência de uma raiz estativa de fonologia ‘*banc*’, em que não há evidência de incidência de causalidade externa ou interna. Tem-se, assim, a raiz *banc*, da qual se derivam *banca*, *banco* e tantos outros constituintes lexicais:

$$banc = \lambda f.(f(r(ve' \langle +st, -in, -ec, -ic \rangle))) \text{ } banc.$$

Tomando-se, como exemplo, a derivação de *banca*, a nominalização acrescenta à raiz três diferentes categorias: (a) a própria categorização da raiz como *nome*; (b) o gênero gramatical, *feminino*; e (c) o número gramatical, *singular*. Analisemos como cada uma dessas categorias é incluída pelo processo de nominalização.

Por princípio, um nome possui semântica estativa, *+st*; a semântica estativa é inerente ao operador *n*, não sendo necessário incluí-la explicitamente entre os traços semânticos da raiz. Neste exemplo particular, a inclusão do traço nada acrescentaria à semântica da raiz, uma vez que a mesma já é estativa.

A nominalização acrescenta à raiz marcas de número e de gênero. Considerarei que, independentemente de motivação semântica (fortemente motivada para número e, na maioria dos casos, puramente convencional para gênero), número e gênero constituem categorias puramente gramaticais, necessárias, em português, à expressão fonológica dos nomes e aos processos de concordância. Embora desvinculadas de qualquer semântica factual (i.e., de base empírica), as categorias gramaticais de número e gênero serão consideradas como pertencentes ao sistema semântico, porém indicadoras de uma semântica puramente gramatical. Deste modo, os valores *masc* e *fem* (gênero) e *sing* e *pl* (número) serão implementados como traços semânticos (não universais) possuidores de expressão fonológica.

Assim, os nomes *banca*, *marca*, *bola* categorizados como femininos e singulares, resultam diretamente da categorização de uma raiz pelo núcleo *N*, em que a semântica, *s*, e a fonologia, *p*, são parametrizadas, respectivamente, por $\langle +sing, +fem \rangle$ e *null*.

$$N \langle +sing, +fem \rangle \text{ } null = \lambda ie.(\lambda f.(f(((n \langle +sing, +fem \rangle \text{ } null)) i)) e).$$

A parametrização da fonologia como *null* implica que o núcleo não agrega uma contribuição fonológica própria, como ocorre com a parametrização denotada por N_{EZA} , anteriormente especificada. Qualquer acréscimo fonológico decorrente do núcleo é puramente gramatical, resultante dos seus traços semânticos.

O núcleo $N \langle +sing, +fem \rangle \text{ } null$ é o formador de nomes categorizados como “feminino,

singular”, em que nada se acrescenta fonologicamente à raiz. Trata-se, portanto, de um núcleo típico, ao qual daremos tratamento lexical por meio das definições

$$N_{SF} = N \langle +sing, +fem \rangle null,$$

e

$$n_{SF} = n \langle +sing, +fem \rangle null.$$

Deste modo, o nome *banc*, a ser pronunciado como 'banca', resulta da raiz *banc*, ao ser categorizada pelo nominalizador N_{SF} . Isto é:

$$banc = banc N_{SF}.$$

Na expressão acima, substituindo a raiz *banc* por sua expressão computacional explícita, tem-se:

$$banc = (\lambda f.(f(r(ve' \langle +st, -in, -ec, -ic \rangle)))) banc) N_{SF},$$

que é equivalente a

$$banc = \lambda f.(f(n_{SF}(r(ve' \langle +st, -in, -ec, -ic \rangle)))) banc.$$

O constituinte acima mostra a composição interna do *nP banc*: uma sequência de traços semânticos e uma de traços fonológicos ($\langle +st, -in, -ec, -ic \rangle$ e *banc*, respectivamente) sob efeito dos operadores *ve'*, *r* e n_{SF} , nesta ordem. O *nP banc* *virtualmente* possui os traços semânticos $\langle +sing, +fem, +st, -in, -ec, -ic \rangle$, resultantes da concatenação dos traços semânticos introduzidos por n_{SF} e pela composição $r \circ ve'$; entretanto os traços semânticos do *nP* somente são efetivamente computados em etapas posteriores da derivação, por operações pós-sintáticas. Um ponto importante, a ser ressaltado, é a composição dos traços semânticos situados na base da derivação, $\langle +st, -in, -ec, -ic \rangle$, proveniente do elemento vocabular categorizado como raiz. Uma raiz dotada de tais traços será denominada *raiz de estatividade pura*. Raízes deste tipo proporcionam uma espécie de base mínima à geração de constituintes *nP*. Embora as raízes não sejam lexicalmente categorizadas, as raízes de estatividade pura podem ser denominadas “raízes nominais”. Nomes são também derivados a partir de outros tipos de raízes, como se verá à continuação.

Capítulo 7

Estruturas estativas

7.1 Raízes estativas básicas

A partir deste momento, as análises terão em vista a derivação verbal, iniciando pelas estruturas estativas mais fundamentais. Exploraremos as possibilidades sintáticas das raízes que estão na base da derivação de estruturas como:

- (32)
- a. A montanha é fria
 - b. O café frio [chegou]
 - c. [Ele foi para o] frio
 - d. O café esfriou
 - e. Maria esfriou o café

Nos exemplos, a raiz *fri* faz parte dos adjetivos *fria/frio*, do nome *frio* e do verbo *esfriou*. Exceto no caso do nome, *frio*, a raiz encontra-se em relação como um nome (*montanha* ou *café*), e a relação sempre caracteriza um certo estado do nome. A raiz *fri*, que dá origem a adjetivos, nomes e verbos é uma raiz estativa (possuidora do traço +*st*), um dos dois tipos de raiz. As estruturas sintáticas a que uma raiz dá origem dependem fundamentalmente dos seus traços semânticos. Iremos, então, determinar os traços semânticos da raiz *fri*, que a tornam compatível com os exemplos. Como vimos, a raiz é detentora do traço +*st*, e é então necessário determinar a composição dos traços *in*, *ec* e *ic*.

Como parte do adjetivo ou do verbo, a raiz participa de relações com outro termo, o que a dota do traço de incidência. A incidência pode ser necessária, +*in*, ou possível (mas não-obrigatória), *in*. Ao formar o nome, *frio*, a raiz não requer complemento, portanto a incidência é não-obrigatória, *in*.

A declaração 'A montanha é fria' expressa uma qualidade própria da montanha, independente de qualquer indício de causalidade (interna ou externa) expressa na própria sentença. Neste caso, semanticamente a raiz não contém requisitos de causalidade, fato que a dota dos traços –*ec* e –*ic*. Embora, nas sentenças verbais com *esfriou*, haja indícios de causalidade, interna

e externa, devem-se buscar os traços semânticos “mínimos” da raiz, a configuração mínima de traços compatível com todos os exemplos.¹

Assim, a raiz *fri* é dotada dos traços semânticos $\langle +st, in, -ec, -ic \rangle$.

Como se trata da determinação de uma configuração mínima de traços, cada exemplo analisado pode manifestar a necessidade de *fortalecimento* dos requisitos da lista, como a inclusão do traço *ic*, no lugar de *-ic*. Mas não pode requerer o *enfraquecimento* de requisitos, como a retirada de *in*, para a entrada de *-in*.

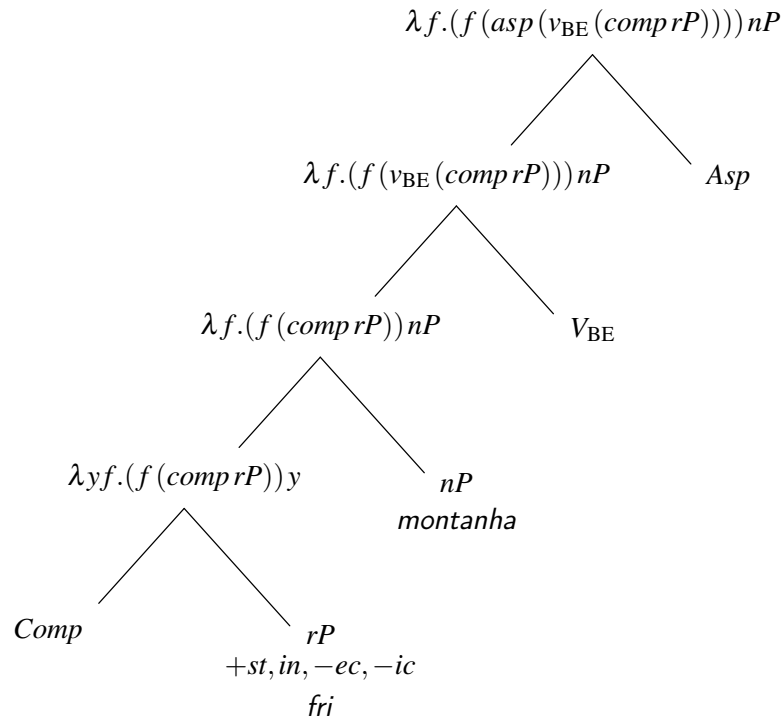
Os traços $\langle +st, in, -ec, -ic \rangle$ traduzem a *possibilidade de uma relação estativa* (traços *+st* e *in*) reduzida aos elementos mínimos, destituída de qualquer referência a *autoria ou situação de iniciação* (traços *-ec* e *-ic*), constituindo a base da predicação de qualidade. As raízes deste tipo dão origem, talvez universalmente, aos adjetivos. Denominarei as raízes deste tipo por “raízes de estatividade básica”. A denominação, de certo modo, é inapropriada, pois a estatividade mais básica é a resultante dos traços $\langle +st, -in, -ec, -ic \rangle$, característica das raízes de estatividade pura. Contudo, como estas raízes não originam diretamente relações estativas, em decorrência do traço *-in*, a denominação pode ser adotada.

7.2 Derivação estativa básica

Passemos à derivação do predicado nominal ‘A montanha é fria’, resultante da verbalização da raiz *fri* pelo núcleo estativo V_{BE} , derivação mostrada no próximo diagrama. O verbalizador estativo V_{BE} , cuja parametrização foi apresentada na seção 5.10, p. 58, se expressa por *default* como os verbos *ser* ou *estar* em português, *to be* em inglês etc.. V_{BE} verbaliza estruturas de complementação portadoras dos traços $\langle +st, in, -ec, -ic \rangle$; a análise das configurações de traços semânticos passíveis de serem verbalizadas por V_{BE} é tema da seção 7.4, p. 95.

¹Os traços de causalidade do verbo *esfriou* resultam de processos alheios à raiz, como veremos. O motivo de se buscarem os traços semânticos mínimos presentes em todos os exemplos é a tentativa de isolar os traços próprios da raiz, separados dos traços acrescentados por outros processos gramaticais.

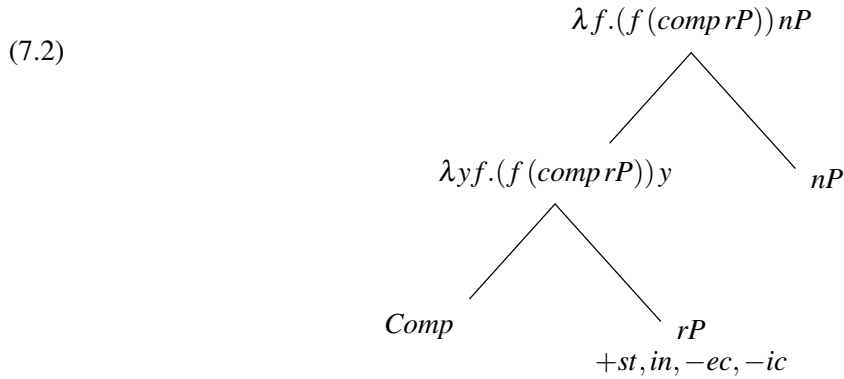
(7.1)



Analisaremos a derivação a partir da raiz,² passo a passo, comentando as operações sintáticas envolvidas. Em seguida abordaremos alguns importantes conceitos que fundamentam a pesquisa como um todo.

(a) *Primeiro passo — relação de complementação*: Como sabemos, a raiz possui fonologia *fri* e traços semânticos $\langle +st, in, -ec, -ic \rangle$. Em decorrência do traço *in*, de incidência, a raiz pode ser posta em relação de complementação com um constituinte *nP* (no caso, *montanha*), o que é feito por meio do núcleo *Comp*. Em outros termos: o núcleo *Comp* é compatível com a raiz e com o *nP*, respectivamente como seus argumentos interno e externo. A aplicação funcional $(ComprP)nP$ produz como resultado o constituinte $\lambda f.(f(comprP))nP$.

²Na realidade, como já foi discutido anteriormente, uma derivação efetivamente se inicia pela formação de um *elemento vocabular*. Entretanto, o segundo passo, necessariamente, é a derivação de uma raiz, sobre a qual se constrói toda a derivação subsequente. Assim, pode-se dizer que uma derivação tem o seu início a partir de uma raiz, sem se perder a generalidade do conceito de derivação.



(b) *Segundo passo — verbalização por V_{BE}* : Os traços semânticos da raiz tornam o constituinte recém-derivado compatível com a verbalização por V_{BE} (ver p. 95), efetuada por meio da aplicação

$$(\lambda f.(f(\text{comp}rP))nP) V_{BE}.$$

Resulta, assim, o constituinte

$$\lambda f.(f(v_{BE}(\text{comp}rP)))nP,$$

um constituinte vP .

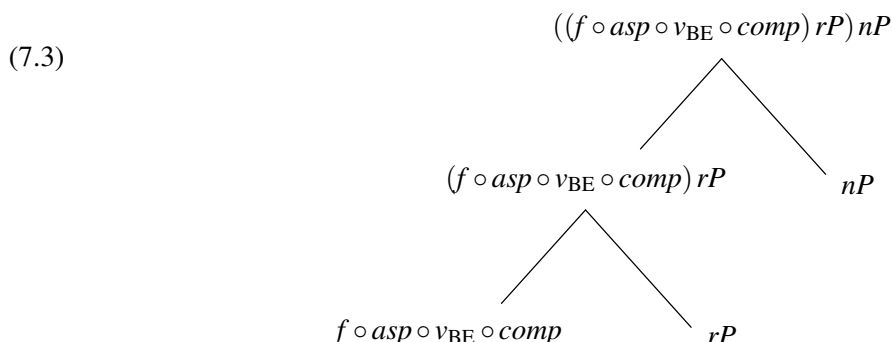
(c) *Terceiro passo — aplicação de aspecto*: Ao se aplicar o núcleo *Asp*, atribuidor de aspecto, ao vP acima formado, deriva-se o constituinte a ser pronunciado como ‘A montanha é fria’, $\lambda f.(f(\text{asp}(v_{BE}(\text{comp}rP))))nP$, exibido no topo do diagrama. Verdadeiramente a derivação ainda não se encontra sintaticamente concluída; é necessária, pelo menos, a inclusão do núcleo *Tense* acima de *Asp*. Este e outros núcleos eventualmente faltantes serão deixados à parte da derivação, por não contribuírem significativamente para as análises estruturais a serem feitas. De todo modo, devemos recordar, todo constituinte encontra-se estruturalmente preparado para absorver operadores provenientes do ambiente, por meio da variável f que o encabeça.

A formação da sequência fonológica a ser pronunciada é obtida por um procedimento, de certo modo, direto:

(a) O constituinte $\lambda f.(f(\text{asp}(v_{BE}(\text{comp}rP))))nP$ possui a estrutura

$$((f \circ \text{asp} \circ v_{BE} \circ \text{comp})rP)nP.$$

Isto é, a estrutura de um operador (neste caso, composto) aplicado binariamente a seus argumentos, de acordo com o padrão canônico:



(b) Nesta estrutura, a raiz, rP , é o argumento interno da cadeia de operadores

$$f \circ asp \circ v_{BE} \circ comp,$$

e o nP é o argumento externo da mesma cadeia. A cadeia de operadores codifica a história derivacional do constituinte; os dois argumentos lexicais, rP e nP , ocupam as mesmas posições estruturais, com relação ao operador, que já ocupavam no início da derivação, em que foram introduzidos pelo núcleo *Comp*.

(c) A cadeia de operadores contextualiza v_{BE} como pertencente a um ambiente de complementação, lexicalmente modalizado por uma raiz especial (no caso, *fri*); a composição funcional/lexical (cadeia de operadores e raiz) “enxerga” o constituinte nominal como o seu argumento interno.

(d) Deste modo, o constituinte nominal, nP , ocupa a posição de argumento externo da cadeia verbal de operadores, $f \circ asp \circ v_{BE} \circ comp$ (a posição “extrema-direita” da cadeia). Esta é a posição estrutural do argumento a ser interpretado como sujeito de um constituinte verbal maximal; o nP será então pronunciado como ‘montanha’, por ação da cadeia de operadores, complementada por outros operadores atribuídos à variável f , deixados à parte da análise.

(e) O paradigma gramatical interno a v_{BE} , em função do contexto funcional em que está inserido (dado pela composição de operadores), e da concordância com nP , expressará a fonologia gramatical ‘é’. Eventualmente a cadeia de operadores poderá expressar uma fonologia que leve em conta também o componente fonológico de rP , o que não ocorre em português, mas que se realiza, por exemplo, nos idiomas em que existe algo como o verbo “*friar*”, significando “*ser frio*”, comum em línguas americanas.

(f) A raiz, que em português possui fonologia em geral defectiva,³ terá sua fonologia com-

³Segundo a terminologia de Hale & Keyser (2002). Constituintes ou operadores possuem fonologia defectiva se a sua expressão fonológica depender da inserção de material fonológico proveniente de constituintes ou operadores externos. Por exemplo, em português a fonologia das raízes é defectiva; em especial, a expressão fonológica dos adjetivos é estabelecida por concordância.

plementada por concordância com o nP , por ação da cadeia de operadores, sendo então pronunciada como 'fria'. Completa-se, assim, a derivação da sentença 'A montanha é fria'.

Nas análises acima, há um importante ponto a ser destacado: ao longo de toda a derivação, o constituinte nP ocupa a posição de argumento externo da cadeia de operadores em formação. Dito de outro modo: independentemente do número de operadores categorizadores e modalizadores, o constituinte nP ocupa sempre a posição “extrema-direita” do constituinte em formação. Reciprocamente, o constituinte rP ocupa a posição de argumento interno, sob as mesmas condições. Tal fato se deve à natureza binária dos núcleos e ao modo como os operadores se inserem em um constituinte, aplicando-se binariamente aos dois termos lá existentes.

Para finalizar, uma característica importante da raiz é que os traços $-ec$ e $-ic$, indicadores de ausência de causalidade (interna ou externa), tornam a estrutura incompatível com o núcleo verbalizador ativo V_{DO} , como se verá quando da análise das estruturas ativas. Este fato impede a geração de sentenças ativas derivadas diretamente deste tipo de raiz, como 'O café "friou"' ou 'A folha "verdará"'. Ou seja, a operação sintática de verbalização das estruturas estativas puras ou básicas por V_{DO} é bloqueada pelos traços semânticos da raiz, especificamente pelos traços de ausência de causalidade.

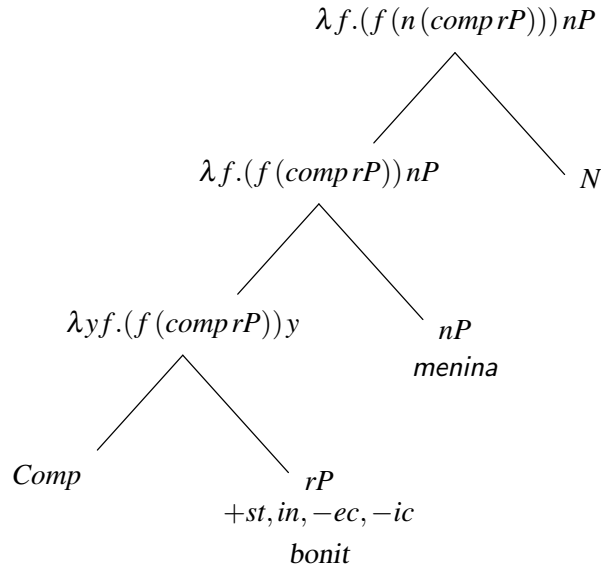
7.3 Derivação nominal

Nas seções 6.5.4 e 6.5.5, analisamos a derivação de nomes a partir de raízes. Abordaremos, agora, a existência de nomes estruturalmente mais complexos.

A derivação da sentença 'A menina é bonita', nos moldes exatos do exemplo anterior, parte da relação de complementação entre dois termos, uma raiz de estatividade básica, *bonit*, e um nP , *menina*. Em 'A menina bonita [chegou]', esses mesmos termos também se encontram em relação de complementação, porém em ambiente nominal, o ambiente nominal também encontrado em '[João cumprimentou a] menina bonita' e '[João deu uma rosa à] menina bonita'. Portanto, tem-se uma relação de complementação utilizada nos mesmos contextos sintáticos em que se utiliza um nome; estruturalmente, tem-se uma relação de complementação nominalizada.

Nos termos da proposta, o constituinte *menina bonita* possui a estrutura de uma relação de complementação categorizada como nome, segundo a derivação a seguir.

(7.4)



A cadeia de operadores do constituinte situado no topo do diagrama mostra que se trata de um constituinte nominal (composição $f \circ n$) em ambiente de complementação (composição $f \circ n \circ \text{comp}$). No constituinte, o único operador dotado de informação categorial é n . Portanto, o constituinte *menina bonita*, da sentença ‘A menina bonita chegou’, possui uma estrutura de nome, ainda que se trate de um nome de estrutura complexa.

Ou seja, o constituinte “inteiro”, *menina bonita*, é um nome complexo, no entanto dotado da propriedade de uma de suas partes, uma raiz (*bonit*), completar a sua fonologia por concordância.

Não se tem, portanto, no constituinte, nenhum elemento que se caracterize particularmente como um adjetivo. A este assunto voltaremos ainda neste capítulo.

É importante observar que a combinação especial de traços das raízes estativas básicas (estatividade, não necessidade de incidência, inexistência de causalidade) possibilita a nominalização imediata da raiz ([o] ‘belo’, [o] ‘azedo’, [o] ‘vermelho’, [o] ‘magro’, [o] ‘raso’, etc.), por meio da derivação

$$rPN,$$

que resulta em um constituinte de estrutura

$$\lambda f.(f(n(r\langle +st, in, -ec, -ic \rangle)))\langle \phi_i \rangle,$$

em que a cadeia de operadores $n \circ r$ nominaliza as sequências de traços semânticos e fonológicos. Deste modo, deriva-se o nome *alto* de ‘Ele foi para o alto’.

As raízes estativas básicas possuem a notável propriedade de poderem ser verbalizadas por V_{BE} , como vimos na seção anterior. Elas mostram agora uma outra importante característica, possibilitarem a geração de estruturas nominais, simples e complexas.

A possibilidade estrutural de nominalização mostra que raízes deste tipo são fundamentalmente destituídas de pressupostos de temporalidade. Mesmo nos idiomas em que o nome contém marcas de tempo ou aspecto (caso, por exemplo, das línguas tupi), pode-se mostrar que tais características resultam de acréscimos semânticos independentes da raiz.

As duas propriedades, categorização por V_{BE} e por N , possuem uma conexão profunda, resultante dos traços $+st$ e in , por um lado, e de $-ec$ e $-ic$, por outro. Este é o tema da próxima seção.

7.4 Verbalização estativa

O fato de as raízes dotadas dos traços semânticos $\langle +st, in, -ec, -ic \rangle$ serem passíveis de nominalização (simples e complexa) proporciona evidências de uma propriedade velada. Nomes (em português e nas línguas indo-europeias, em geral) são gramaticalmente destituídos de semântica temporal. Qualquer referência temporal é claramente associada a uma atribuição bem caracterizada, como em *nascituro* ou em *futura-esposa*.

O ponto fundamental é que V_{BE} verbaliza uma raiz destituída de semântica temporal, e que a ela nada acrescenta além do caráter verbal (cada núcleo introduz apenas um efeito). Portanto, o predicado nominal é puramente estativo, embora de caráter verbal; os traços semânticos o mostram. As marcas de aspecto e tempo (*é*, *era*, *será*) resultam de acréscimos por outros núcleos, independentes de V_{BE} .

Farei, então, a hipótese altamente restritiva de que o verbalizador V_{BE} é compatível apenas com as raízes de estatividade básica, nas situações em que a raiz se encontra em relação de complementação com um nP ; isto é, V_{BE} deriva diretamente (sem a interveniência de outros núcleos) apenas predicados nominais.

A hipótese significa que a verbalização estativa somente pode ocorrer em situações destituídas de causalidade. Há um pressuposto implícito nesta decisão: os traços de causalidade, *ec* e *ic*, comprometem a raiz com algum tipo de processualidade, tornado-as não-estativas.

Resulta daí uma consequência igualmente fundamental: as raízes dotadas de traços de causalidade semanticamente denotam processos. Ou seja, não apenas raízes dotadas de traço $+ac$ denotam processos, mesmo raízes estativas, desde que dotadas de traços de causalidade, são também processuais.

Raízes estativas portadoras de traços de causalidade originam, por exemplo, os verbos inacusativos, tipicamente verbos estativos, como *chegar* e *morrer*. Segundo o modelo que proponho, tais raízes somente podem ser verbalizadas pelo núcleo V_{DO} , consideradas as restrições impostas a V_{BE} .

A verbalização por V_{DO} desta classe de raízes estativas explica o motivo de, em muitas famílias linguísticas, os sujeitos de verbos inacusativos, mesmo não exibindo semântica ativa, receberem o mesmo tratamento gramatical aplicado aos sujeitos de verbos “de ação”, como os

transitivos e alguns denominais. Tal tratamento dos sujeitos caracteriza o sistema nominativo-acusativo de atribuição de caso, existente em latim, português e inglês, por exemplo.

Segundo a presente proposta, a verbalização por V_{BE} somente se aplica a uma classe muito restrita de raízes, aquelas dotadas dos traços semânticos

$$\langle +st, in, -ec, -ic \rangle \text{ ou } \langle +st, +in, -ec, -ic \rangle;$$

además, as raízes dotadas dessas configurações de traços somente podem ser verbalizadas por V_{BE} . A verbalização por V_{DO} independe dos traços de estatividade ou atividade da raiz, a única condição é a inexistência conjunta dos traços $-ec$ e $-ic$.

Capítulo 8

Estruturas ativas

8.1 Estrutura inacusativa

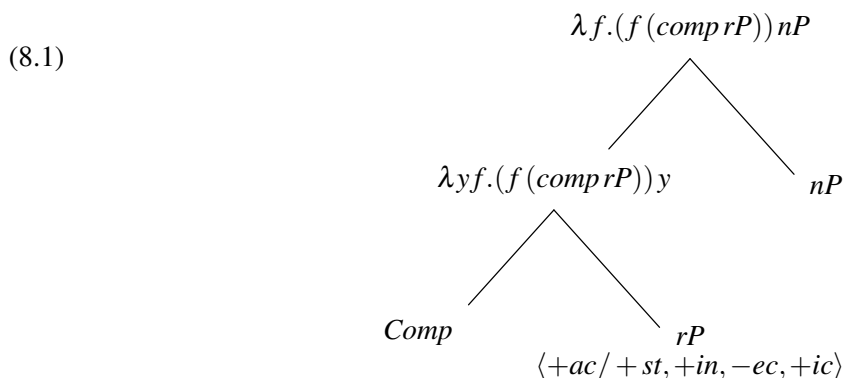
Nos anos 70, a Hipótese Inacusativa de Perlmutter (1978) deu uma grande contribuição à teoria da gramática, reconhecendo a existência de dois tipos de verbos intransitivos: os *inacusativos*, em que o sujeito tem origem no argumento interno do verbo; e os *inergativos*, em que a origem do sujeito é o argumento externo do verbo. Até a formulação da Hipótese, no âmbito gerativo, todos os sujeitos dos verbos intransitivos eram gerados no componente de base a partir de regras como

$$(57)(i) S \rightarrow NP \frown \text{Predicate-Phrase},$$

de Chomsky (1965a).

A investigação das estruturas inacusativas foi levada adiante por Burzio (1986), na Teoria da Regência e Ligação, em bases eminentemente sintáticas. Burzio enfatizou que todos os verbos inacusativos compartilham determinadas características sintáticas, como a seleção de um argumento interno e a impossibilidade de atribuir caso acusativo a este argumento, além da incompatibilidade com argumentos externos. Burzio também estabeleceu critérios de caracterização de inacusatividade para a língua italiana, como: (a) a seleção do auxiliar *essere* para a formação do tempo passado; (b) a concordância do particípio com o seu sujeito; e (c) a extração do objeto direto por meio do clítico *ne*.

No modelo que proponho, a diferenciação entre verbos inacusativos e inergativos é expressa sintaticamente, mas é semanticamente codificada em propriedades lexicais da raiz. Para se falar com propriedade, no modelo, os “verbos” inacusativos e inergativos são gerados a partir da *verbalização* de “estruturas” inacusativas e inergativas. A *estrutura inacusativa* caracteriza-se por uma raiz dotada dos traços semânticos $\langle +ac / +st, +in, -ec, +ic \rangle$, posta em relação de complementação com um *nP*, conforme a representação do diagrama a seguir.

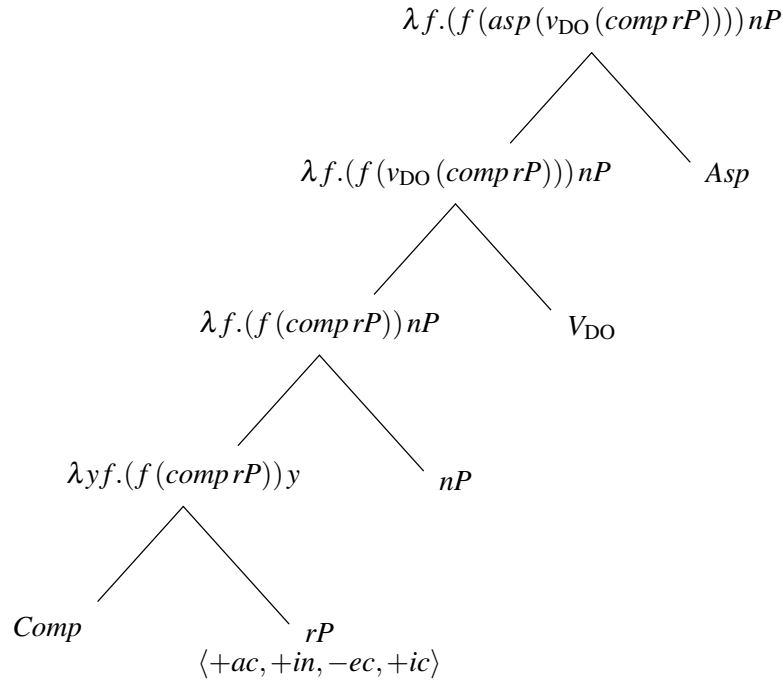


A seleção dos traços semânticos decorre de vários fatos estruturais. O par de sentenças (34) '[Eu vi] a carta caída'/(35)'A carta está caída' mostra estruturas estativas (fora dos colchetes). Porém a estatividade das sentenças não é derivada diretamente da raiz, é formada por meio de um processo suplementar, um particípio. Trata-se, assim, de uma raiz não-estativa, sendo, portanto, dotada do traço de atividade $+ac$. A formação do particípio, em 'carta caída', também mostra que o nP , 'carta', foi introduzido na derivação como complemento da raiz de 'caída', fato que comprova o traço de incidência $+in$, uma vez que a incidência é necessária a todos os exemplos. A agramaticalidade de (36) 'Maria caiu a carta' demonstra a incompatibilidade lexical da raiz com a causalidade externa, daí resultando o traço $-ec$. A sentença (33)'A carta caiu' denota a compatibilidade lexical da raiz com a causalidade interna, o que proporciona o traço ec ; como está excluída a causalidade externa, tem-se então o traço $+ic$.

- (33) A carta caiu
 (34) [Eu vi] a carta caída
 (35) A carta está caída
 (36) *Maria caiu a carta

Analisemos as implicações estruturais dos traços semânticos: (a) o traço de incidência $+in$ torna a raiz compatível com o núcleo *Comp* de complementação; (b) o traço semântico $+ic$ possibilita à estrutura de complementação ser verbalizada por V_{DO} e exclui a verbalização estativa imediata por V_{BE} . Tem-se então a derivação a seguir, em que também foi acrescentado o núcleo *Asp*, necessário à formação do particípio.

(8.2)



O traço $-ic$ da raiz, que denota a ausência necessária de causalidade externa, elimina a possibilidade de um núcleo específico acrescentar um segundo constituinte nP à estrutura situada no topo da derivação,¹

$$\lambda f.(f(asp(v_{DO}(comp\ rP))))nP,$$

nP que normalmente tornar-se-ia o sujeito de uma sentença transitiva. Exclui-se, assim, a geração da sentença transitiva (36) ‘Maria caiu a carta’. Deste modo, a estrutura acima encontra-se finalizada, no que diz respeito à composição argumental. Por facilidade de exposição, esta estrutura será denominada *estrutura inacusativa*, embora a denominação mais precisa seja *estrutura inacusativa verbalizada*, uma vez que a inacusatividade se deve às propriedades lexicais da raiz e não ao processo de verbalização.

8.2 Noção de voz ativa

Para a derivação de uma sentença, como ‘A carta caiu’, a partir do constituinte

$$\lambda f.(f(asp(v_{DO}(comp\ rP))))nP,$$

situado no topo do diagrama anterior, é necessário compreender o modo como o nP situado à extrema-direita do constituinte se torna sujeito sentencial.

¹O núcleo que acrescenta argumentos externos a uma estrutura verbal é *Voice*, introduzido por Kratzer (1996)., ao qual voltaremos na p. 106 ao analisar a derivação das estruturas transitivas.

Em primeiro lugar, deve-se ter em conta que o constituinte derivado *ainda* não é especificamente verbal. Dependendo do valor atribuído à variável f , a associação do nP com $(asp(v_{DO}(comp rP)))$ pode mesmo receber um tratamento nominal: basta que se atribua a f o valor N , o núcleo nominalizador, caso asp tenha sido previamente definido como participio; deriva-se então o nome ‘carta caída’.

Portanto, o núcleo Asp deixa em aberto a possibilidade de a raiz verbalizada receber um tratamento de adjetivo verbal (ou seja, de participio), permitindo à estrutura ser nominalizada. Estruturalmente, isto significa que a relação entre os dois argumentos da variável f não se encontra ainda fixada (no caso, f possui o argumento interno $(asp(v_{DO}(comp rP)))$ e o argumento externo nP).

O carácter especificamente verbal do relacionamento entre os dois argumentos de f é dado pelo núcleo especificador de voz, o núcleo Vox .² Trataremos da noção de voz em capítulo vindouro, a partir de uma proposta de Benveniste (1950a); entretanto, é necessário introduzir o assunto neste ponto, para que se possa caracterizar adequadamente a estrutura do constituinte verbal.

O núcleo Vox caracteriza a relação entre os dois argumentos do constituinte verbalizado, nP e $(asp(v_{DO}(comp rP)))$, comprometendo esta relação com uma interpretação verbal. A caracterização específica da relação depende da semântica da raiz, que licencia, de modo não exclusivo, uma das três vozes – *ativa*, *média* ou *passiva*.

Como vimos anteriormente, a posição estrutural do sujeito é a de argumento externo de um constituinte verbalizado. Assim, a voz qualifica o tipo de relação existente entre o sujeito e o argumento verbal.

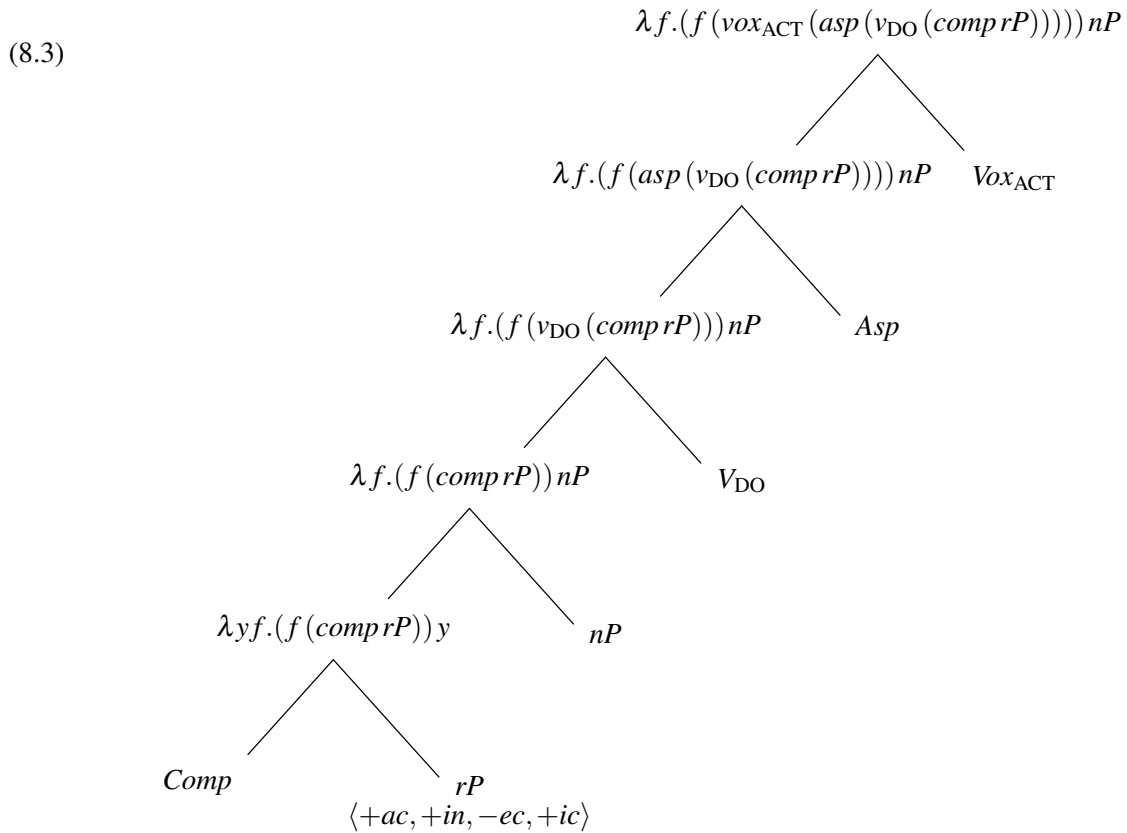
A semântica fundamental da voz ativa caracteriza-se pelo fato de o sujeito ser necessário à iniciação do processo ou estado indicado pelo verbo.³ Em sentenças verbalizadas por V_{DO} , o argumento externo introduzido em decorrência dos traços *ec* ou *ic* enquadra-se nesta categoria. É importante observar que os traços de causalidade licenciadores da voz ativa, no caso de verbalização por V_{DO} , podem não provir da raiz; como veremos, podem ser introduzidos por preposições.

8.3 Derivação verbal inacusativa

A derivação da sentença ‘A carta caiu’ é mostrada no diagrama a seguir:

²O termo *Voice* já se encontra consagrado à introdução de argumentos externos a uma estrutura verbal; por este motivo, utilizo a denominação *Vox* para o núcleo especificador de voz.

³Na voz passiva, ‘A bola foi/será chutada’, e na média, ‘A casa incendiou-se’, a iniciação do evento indicado pelo verbo fundamentalmente independe do sujeito; qualquer dependência existente resulta de fatores colaterais, como a agentividade lexical do sujeito.



Faremos uma análise da derivação do constituinte situado no topo do diagrama: este constituinte é a estrutura sintática e semântica de uma sentença inacusativa básica, excluídos os núcleos situados acima de vox_{ACT} . Não foi considerado, por exemplo, o núcleo *Tense*, por não contribuir para os fenômenos de formação da estrutura argumental; de todo modo, *Tense* se “infiltra” na estrutura por meio da variável f .

- (a) A raiz possui os traços $\langle +ac, +in, -ec, +ic \rangle$;
- (b) O traço $+in$ licencia a inclusão do nP , por meio do núcleo *Comp*;
- (c) O traço $+ic$ licencia a verbalização por V_{DO} , e a subsequente inclusão de *Asp*;
- (d) E a estrutura resultante é compatível com a voz ativa, vox_{ACT} , fundamentalmente em decorrência do traço $[+ic]$ da raiz, que licencia a interpretação do nP como sujeito estrutural;
- (e) O verbo é formado pela relação entre a cadeia de operadores

$$f \circ vox_{ACT} \circ asp \circ v_{DO} \circ comp$$

e o argumento interno da cadeia, a raiz rP . (O verbo efetivamente “é” esta relação; a estrutura verbal é assim formada inteiramente na sintaxe.) Todo operador é interpretado contextualmente, de acordo com os demais operadores da cadeia de que participa, situados à sua esquerda e di-

reita. A análise da formação da fonologia do verbo deve tomar por base o operador verbalizador, no caso v_{DO} . No contexto em que se encontra, v_{DO} adquire fonologia “apropriando-se” do componente fonológico de rP .⁴ No caso do verbo *cair* (de ‘A carta caiu’, a sentença em análise), suporei que a raiz possui fonologia $ca[d]$. Além de absorver a fonologia da raiz, o verbalizador acrescenta sua própria contribuição fonológica; em português, o acréscimo é a vogal temática. Assim, a fonologia parcialmente expressa por v_{DO} é *cai* ($ca + i$). Neste ponto, a fonologia do verbo deve expressar também as contribuições fonológicas dos operadores *asp* e vox_{ACT} , situados à esquerda de v_{DO} na cadeia de operadores. Em português, no exemplo em consideração, ambas as contribuições são nulas.

(f) Deve-se ter em conta, entretanto, que todos os processos morfossintáticos do constituinte $\lambda f.(f(vox_{ACT}(asp(v_{DO}(comprP))))nP$, situado no topo do diagrama, encontram-se em estado suspenso; a variável f interrompe todas as operações presentes em seus dois argumentos. Assim, o nP é um sujeito estrutural, como analisaremos na próxima seção, mas não exhibe, no constituinte, qualquer expressão morfossintática de caso. Também a morfossintaxe verbal encontra-se em estado “virtual”. Este fenômeno é resultante do *processamento postergado* inerente ao meu modelo, o que torna desnecessária a postulação da inserção tardia (*late insertion*) da fonologia, um princípio fundamental da Morfologia Distribuída. Somente após a aplicação dos núcleos situados acima de Vox_{ACT} a sentença poderá ser pronunciada. Isto inclui núcleos como *Tense*, que contribui para a morfossintaxe verbal, e outros núcleos responsáveis pelas operações de atribuição de caso e concordância. Provavelmente as operações que compõem o constituinte se completarão apenas quando da execução da última operação pós-sintática, quando então a sentença será pronunciada.

8.3.1 O sujeito inacusativo

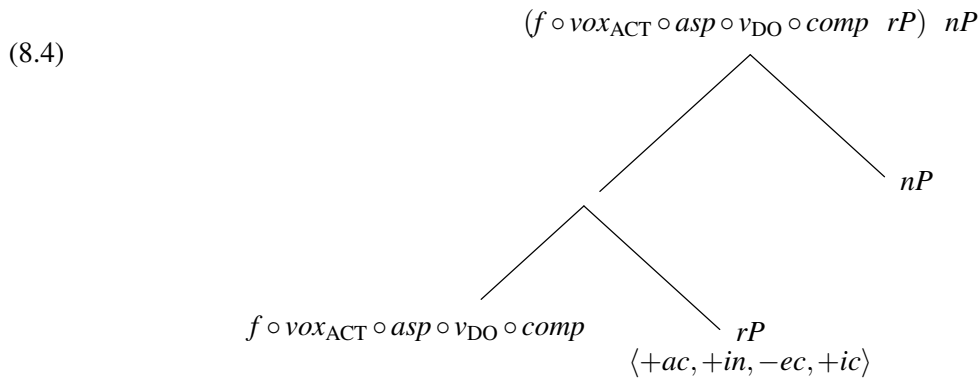
Analisemos a caracterização do sujeito da estrutura inacusativa, o constituinte nP situado na posição extrema-direita do constituinte maximal

$$\lambda f.(f(vox_{ACT}(asp(v_{DO}(comprP))))nP,$$

resultante da derivação mostrada no diagrama anterior.

As propriedades estruturais do constituinte podem ser melhor observadas se a estrutura for reinterpretada em termos de composição de núcleos, de acordo com o diagrama abaixo.

⁴Este é o processo de *conflation* proposto por Hale & Keyser (2002). Creio que o processo possui realidade linguística, não se constituindo em mero recurso formal. Há diferenças entre o modo como o mecanismo de *conflation* é implementado no meu modelo e na proposta de H&K, porém o conceito fundamental se deve inteiramente aos dois autores.



O diagrama apresenta uma estrutura binária, em que o operador verbal composto

$$f \circ \text{vox}_{\text{ACT}} \circ \text{asp} \circ \text{v}_{\text{DO}} \circ \text{comp}$$

tem a raiz, rP , como argumento interno e o constituinte nominal, nP , como argumento externo. A posição estrutural do sujeito inacusativo é a de argumento externo do operador verbal composto (como é mostrado no diagrama), em que vox_{ACT} licencia o nP como sujeito estrutural, em decorrência do traço $[+]ic$ da raiz. (Como veremos na próxima seção, o traço $-ec$ exclui o acréscimo de um segundo nP à estrutura já formada, nP que viria a ser o sujeito de uma sentença transitiva.) A cadeia de operadores caracteriza o nP como sujeito estrutural, atribuindo-lhe uma expressão de caso (morfológica ou sintática) por processos pós-sintáticos.

De modo geral, a posição do sujeito de uma estrutura verbal é a de argumento externo do constituinte verbal maximal, aquele categorizado por v_{DO} ou v_{BE} .

No caso da estrutura inacusativa, uma associação entre nP e rP , que se iniciou com uma relação de complementação, progressivamente foi-se enriquecendo pela anexação de novos operadores, mas sem perder o caráter original de uma associação entre nP e rP .

É importante observar que o nP é (futuro) sujeito (pronunciado) de uma sentença ativa, mesmo tendo sido incluído na derivação como complemento da raiz e abaixo do núcleo verbalizador, V_{DO} , ou seja, mesmo tratando-se de um “argumento interno”. O nP ocupa a posição estrutural de sujeito sem ter havido movimento para uma posição diferente daquela em que foi inicialmente incluído na derivação.⁵

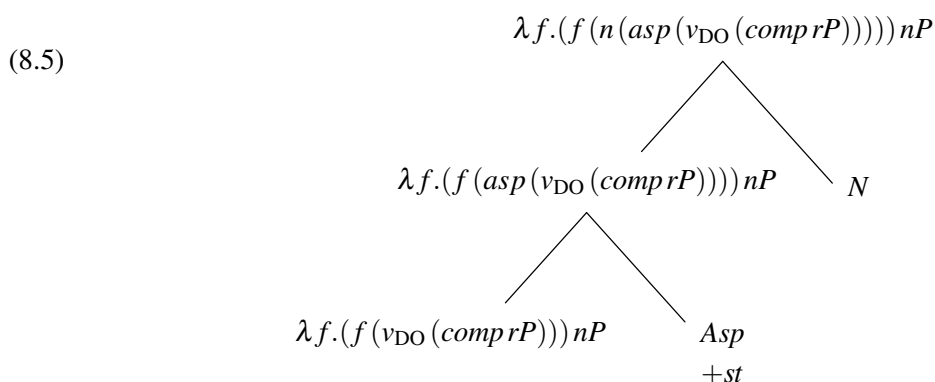
Conforme o analisado na seção anterior, a relação entre o operador composto e a raiz resulta na formação fonológica do verbo, processo comum a todas as estruturas verbais, como veremos em capítulos vindouros. Isto significa que a raiz cumpre funções de formação do verbo e é desvinculada da interpretação como objeto. Portanto, a estrutura inacusativa não possui a posição estrutural de objeto.

⁵No modelo que proponho, não existe a criação de uma posição de especificador verbal, para onde se move o complemento de um verbo para receber caso como sujeito, a solução-padrão da GG.

8.3.2 Estatividade e particípio

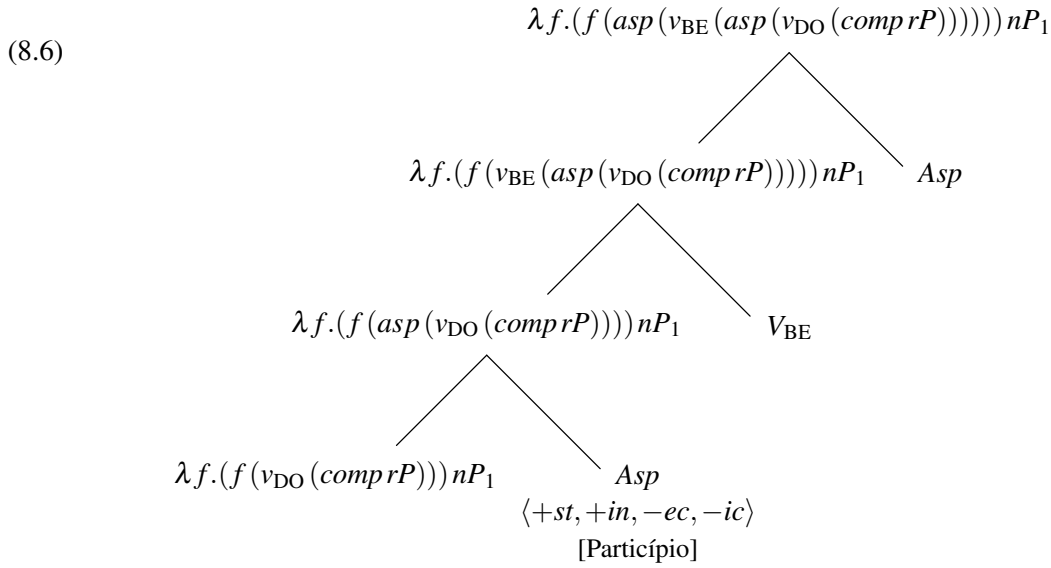
É importante observar que, em um verbo como *cair*, qualquer traço de estatividade independe da raiz e é, necessariamente, um acréscimo à formação do verbo, um processo situado acima de V_{DO} . Este é o motivo de a estatividade dos exemplos ser obtida de um particípio. O particípio, uma realização do núcleo *Asp*, é dotado do traço $+st$, que é acrescentado aos traços da estrutura verbal por composição de operadores.

A estrutura inacusativa em que *Asp* assume o valor de particípio pode ser nominalizada, como é ilustrado no diagrama abaixo. Tem-se, então a derivação de ‘carta caída’, como em (34) ‘[Eu vi] a carta caída’. Neste caso, o operador composto $asp(v_{DO}(comp rP))$ é dotado dos traços de estatividade ($+st$) e incidência ($+in$) característico da formação dos adjetivos, o que permite a nominalização.



Deve ser ressaltado que na estrutura nominalizada, acima, não existe algo como “um verbo *ser/estar* implícito, mudo”, a produzir uma sentença do gênero ‘A carta [está] caída’. Tem-se, efetivamente, a estrutura de um adjetivo verbal, resultante da nominalização do operador composto $asp \circ v_{DO} \circ comp$, em que *asp* e *comp* acrescentam, respectivamente, os traços $+st$ e $+in$, característicos das estruturas estativas, e v_{DO} contribui com a categorização verbal.

A derivação de (35) ‘A carta está caída’ resulta da verbalização, por v_{BE} , da estrutura estativa, como mostrado no diagrama a seguir. Somente é possível a verbalização por v_{BE} graças à introdução do traço $+st$ pelo particípio.



8.4 Verbos inergativos

Na seção anterior, analisamos a estrutura verbal inacusativa, em que o único argumento nominal da estrutura é introduzido na derivação em consequência do traço de incidência da raiz, abaixo do núcleo verbalizador, tornando-se o sujeito estrutural em virtude do traço de causalidade interna. Analisaremos agora a estrutura inergativa, em que o único argumento nominal é introduzido na estrutura acima do núcleo verbalizador.

Uma sentença inergativa típica é:

(37) a. Maria riu

Assumirei que a sentença é derivada a partir da raiz *ri*, subjacente ao verbo *rir*, e tomarei como suficientemente bem estabelecido pela teoria linguística que os verbos inergativos possuem um *argumento externo*. Assim, considerando-se a sentença em sua forma intransitiva mais simples, como a acima, a raiz é dotada dos traços

$$\langle +ac, -in, +ec, -ic \rangle.$$

Esses traços semânticos caracterizam a estrutura inergativa “pura”, destituída de qualquer possibilidade de introdução de um argumento interno. Não há como gerar, por exemplo, uma sentença como ‘Maria riu um sorriso sarcástico’. Os traços semânticos da raiz, de acordo com a suposição, são compatíveis com ‘Maria dança’, mas não com ‘Maria dança tango’. Esta dupla possibilidade, de produção de uma sentença transitiva e de outra intransitiva, é resultante dos traços semânticos $\langle +ac, in, +ec, -ic \rangle$, em que a raiz não requer um argumento interno, embora

tampouco o proíba, mas rejeita a causalidade interna, requerendo a causalidade externa. Provavelmente não há raízes inergativas puras em português, sendo sempre possível a introdução de um argumento semanticamente ligado, de algum modo, ao significado enciclopédico da raiz. O fenômeno de “verbos inergativos” possuírem uma variante transitiva é universal, e analisado em Hale & Keyser (2002).

Existe uma possibilidade que não será explorada neste trabalho. Uma sentença como ‘Maria riu um sorriso sarcástico’ pode ser parafraseada por ‘Maria riu sarcasticamente’. Ou seja, o objeto da variante transitiva de um verbo inergativo daria conteúdo a uma modalização verbal, não se constituindo em um “verdadeiro objeto”, o que explicaria a ligação semântica entre verbo e objeto, nesta classe de verbos. Estrutura similar seria válida também para ‘Maria dança tango’, mas neste caso a modalização da estrutura verbal por um *nP*, como *tango*, não se expressaria de modo “sintético” em português, por meio de algo semelhante a um advérbio: um núcleo modalizador não seria capaz de incorporar o material fonológico de um *nP* ao de um *rP*. Tal interpretação preserva o caráter de um verbo inergativo, como “selecionador de um único argumento, externo”. Entretanto, há muitas implicações, sobretudo com relação ao fenômeno da transitividade, dependentes de pesquisa mais ampla, o que não poderá ser feito nos limites atuais da pesquisa.

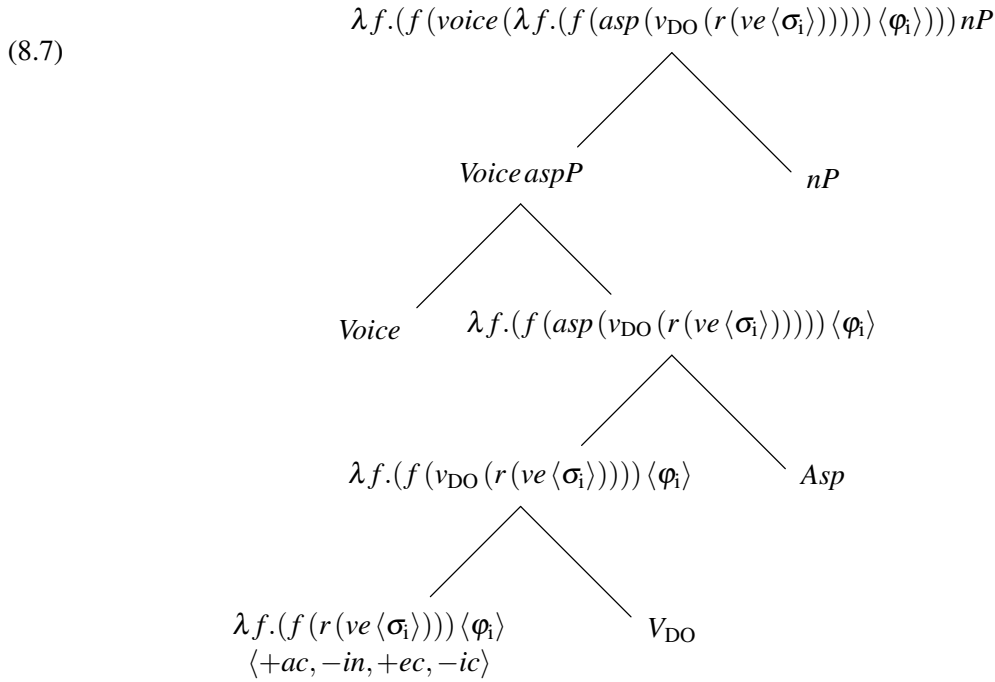
De todo modo, se considerarmos as raízes de atividade que requerem causalidade externa e rejeitam causalidade interna, diferindo apenas no traço de incidência, pode haver: (a) raízes dotadas dos traços $\langle +ac, -in, +ec, -ic \rangle$, caracterizadoras da inergatividade pura; (b) raízes possuidoras dos traços $\langle +ac, in, +ec, -ic \rangle$, que combinam características inergativas e transitivas (raízes inergativas alternantes); e (c) raízes que exibem os traços $\langle +ac, +in, +ec, -ic \rangle$, originadoras da transitividade pura. As configurações de traços dos três tipos de raízes são independentes e, em consequência, as estruturas sintáticas derivadas de cada tipo devem ser independentemente consideradas.

Analisaremos, nesta seção, a estrutura intransitiva gerada a partir das raízes dotadas de traços de inergatividade pura; na próxima seção analisaremos a estrutura transitiva “pura”; e, em seguida, consideraremos a estrutura dos verbos inergativos dotados de “transitividade alternante”.

Os traços $\langle +ac, -in, +ec, -ic \rangle$ dão origem à derivação básica de um verbo inergativo, representada no próximo diagrama.

Os verbos inergativos e transitivos requerem a inserção de um argumento nominal externo ao núcleo verbalizador ativo, isto é, que entre na derivação após V_{DO} . Segundo a presente proposta, argumentos deste tipo são licenciados pelos traços de causalidade externa da raiz, *ec* ou *+ic*. Como toda operação sintática, a inserção do argumento externo é feita por um núcleo especial; por tradição, designarei este núcleo por *Voice*, seguindo a proposta de Kratzer (1996).⁶

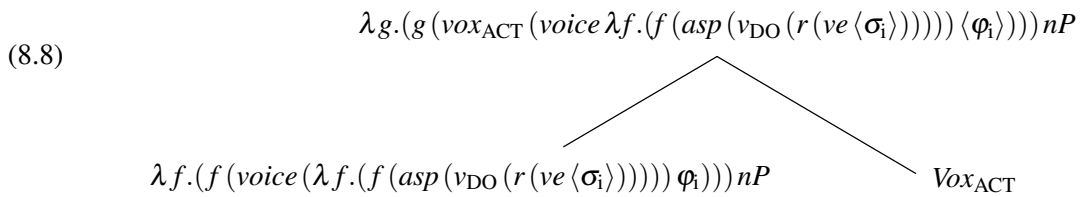
⁶O núcleo *Voice* é um núcleo relacionador, associando um *nP* a um constituinte *aspP*; deste modo é incluído à direita em uma derivação.



- (a) O traço *–in* torna a raiz incompatível com o núcleo *Comp*, fato que impede a introdução do constituinte nominal que viria a tornar-se argumento interno do verbo.
- (b) Portanto, a raiz é diretamente verbalizada por V_{DO} , núcleo licenciado pelo traço $[+ec]$, e modalizada por *Asp*, núcleo licenciado por V_{DO} ;
- (c) O núcleo *Voice*, licenciado pelo traço $[+ec]$, introduz o argumento *nP*. Resulta, assim, a estrutura do topo do diagrama:

$$\lambda f.(f(\text{voice}(\lambda f.(f(\text{asp}(v_{\text{DO}}(r(\text{ve}' \langle \sigma_i \rangle)))))) \langle \varphi_i \rangle)) nP;$$

- (d) Nesta estrutura, o *nP* é então qualificado como iniciador do evento indicado pelo constituinte verbal, por meio do núcleo Vox_{ACT} , licenciado pelo traço $[+ec]$. Isto é, o *nP* é caracterizado como sujeito estrutural pela voz ativa, segundo a derivação mostrada no próximo diagrama.



Gera-se, assim, o constituinte:⁷

$$\lambda g.(g(vox_{ACT}(voice \lambda f.(f(asp(v_{DO}(r(ve' \langle \sigma_i \rangle)))))) \langle \phi_i \rangle)) nP,$$

que é a estrutura verbal inergativa: dados os traços semânticos, $\langle \sigma_i \rangle$, fonológicos, $\langle \phi_i \rangle$, e o nP argumento externo, geram-se sentenças como ‘Maria riu’, após a inclusão dos núcleos situados acima de Vox_{ACT} , omitidos da derivação.

8.5 Verbos transitivos

Analisaremos agora a estrutura dos verbos transitivos, que derivam expressões como:

- (38)
- a. Maria lava o prato/Maria ama a criança
 - b. *O prato lava
 - c. Prato lavado
 - d. O prato está lavado

Se assumirmos o verbo *amar* como estativo (indicador de um *estado do sujeito*), os verbos transitivos possuem os traços semânticos

$$\langle +ac / +st, +in, +ec, -ic \rangle.$$

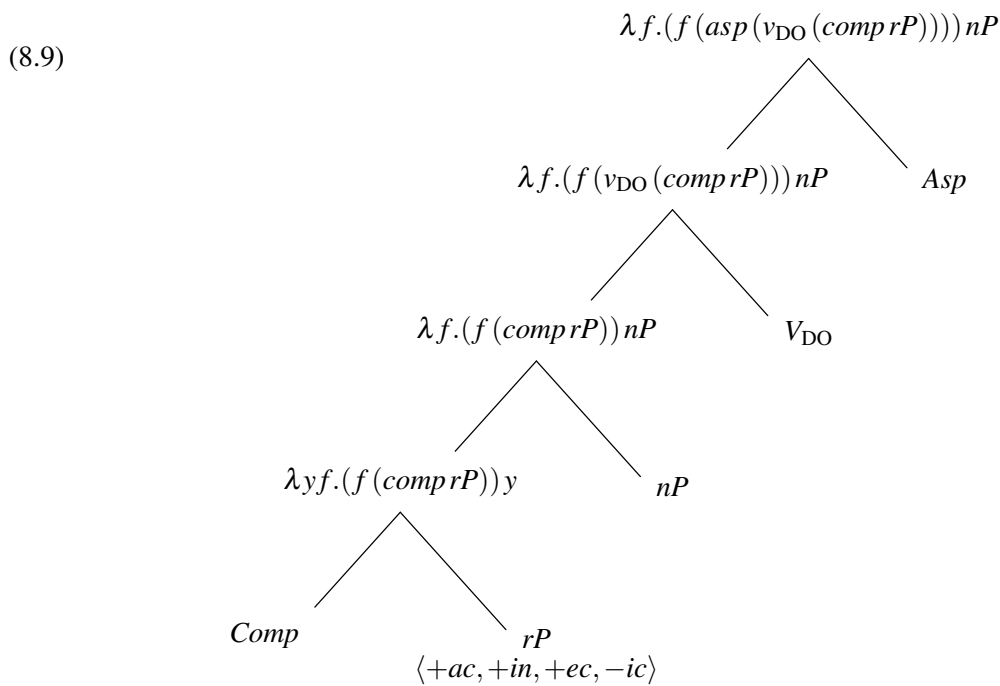
As análises desta seção têm em vista apenas aqueles que se poderiam denominar verbos “transitivos puros”, dotados dos traços semânticos citados acima, e que, como veremos, exibem expressão fonológica adquirida diretamente da raiz, como *lavar* e *amar*. Ficam portanto excluídos os verbos como *quebrar*, que possuem uma variante transitiva (‘João quebrou o rádio’) e outra intransitiva (‘O rádio quebrou’); a alternância de transitividade se deve à coexistência, na estrutura do verbo, dos traços *ec* e *ic*. Também não serão considerados os verbos denominais (‘João assustou Maria’, ‘João engavetou o livro’), por apresentarem fonologia derivada de um nome (‘susto’, ‘gaveta’), e não de uma raiz. Ademais, não farão parte das análises da seção verbos gramaticais como *ter* e *dar*, cuja estrutura e fonologia independem inteiramente de raízes. Os verbos excluídos das análises farão parte de estudos posteriores neste trabalho.

O objetivo da seção é assim enfocar a transitividade resultante exclusivamente dos traços semânticos da raiz, excluindo os efeitos de transitividade provocados por determinados núcleos. Alguns verbos transitivos alternantes, como *quebrar*, também parecem apresentar transitividade lexical, porém muitos verbos alternantes, como *azedar*, dão mostras de possuir transitividade derivada de processos gramaticais independentes da raiz. Por este motivo, os ver-

⁷Relembrar que é irrelevante o identificador utilizado como variável do operador λ . Na derivação, foram utilizados g e f por motivo de clareza; poder-se-iam empregar dois f s, sem ambiguidade, uma vez que cada f possui o seu escopo de atuação.

bos alternantes serão tratados como uma classe à parte, a dos verbos incoativos; possivelmente a incoatividade (*iniciação de um processo*, ou, alternativamente, *mudança de estado*) seja um fenômeno interpretativo resultante de um efeito colateral, como se analisará posteriormente na p. 126.

Analisemos, então, a estrutura dos exemplos apresentados no início da seção, derivados da raiz *lav*, do verbo *lavar*. Trata-se de uma raiz de atividade dotada dos traços semânticos $\langle +ac, +in, +ec, -ic \rangle$; esta configuração de traços dá origem à derivação:



- (a) O traço $+in$ da raiz licencia a inclusão do nP por meio do núcleo *Comp*;
- (b) O traço $+ec$ licencia a verbalização por V_{DO} e a subsequente inclusão de *Asp*;
- (c) Deriva-se, assim, o constituinte $\lambda f.(f(asp(v_{DO}(comprP))))nP$. O traço $-ic$ torna o constituinte incompatível com a voz ativa, uma vez que o nP não pode ser caracterizado como iniciador do processo indicado pelo verbo; entretanto, o constituinte é compatível com as vozes passiva e média, como veremos mais adiante. Devido à incompatibilidade com Vox_{ACT} , a estrutura não pode derivar 'O prato lavou', dada a impossibilidade de verbalização ativa.
- (d) Contudo, caso o operador *asp* se apresente com um valor de particípio, a estrutura pode ser nominalizada, derivando então 'Prato lavado', ou verbalizada por V_{BE} , gerando 'O prato está lavado', como é mostrado nos diagramas a seguir.

(8.10)

$$\begin{array}{c}
 \lambda f.(f(n(asp(v_{DO}(comp rP)))))nP \\
 \swarrow \quad \searrow \\
 \lambda f.(f(asp(v_{DO}(comp rP)))))nP \quad N
 \end{array}$$

(8.11)

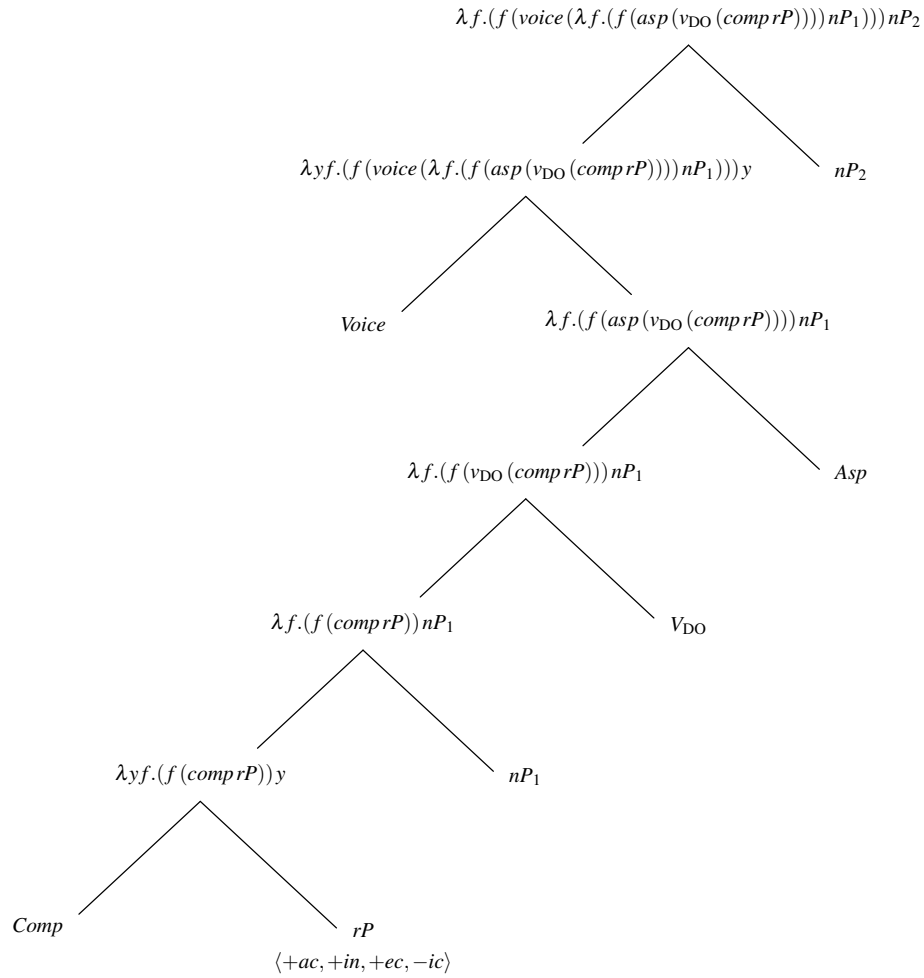
$$\begin{array}{c}
 \lambda f.(f(v_{BE}(asp(v_{DO}(comp rP)))))nP \\
 \swarrow \quad \searrow \\
 \lambda f.(f(asp(v_{DO}(comp rP)))))nP \quad V_{BE}
 \end{array}$$

Há ainda uma possibilidade derivacional proporcionada pelo constituinte

$$\lambda f.(f(asp(v_{DO}(comp rP)))))nP.$$

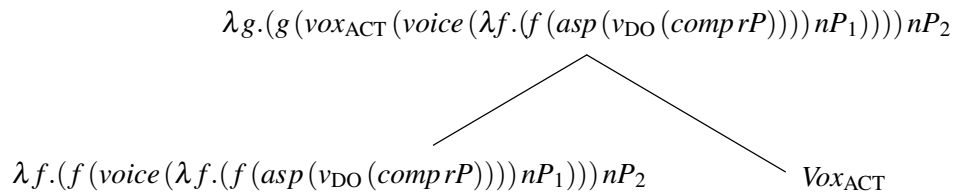
O traço *+ec* torna-o compatível com um argumento independente do traço de incidência da raiz; isto é, o constituinte é compatível com um argumento externo, não requerido pelo traço *in* da raiz. Tem-se, então, a derivação exibida no próximo diagrama (em que os constituintes nominais portam índices meramente diferenciais, isentos de qualquer interpretação de referencialidade).

(8.12)



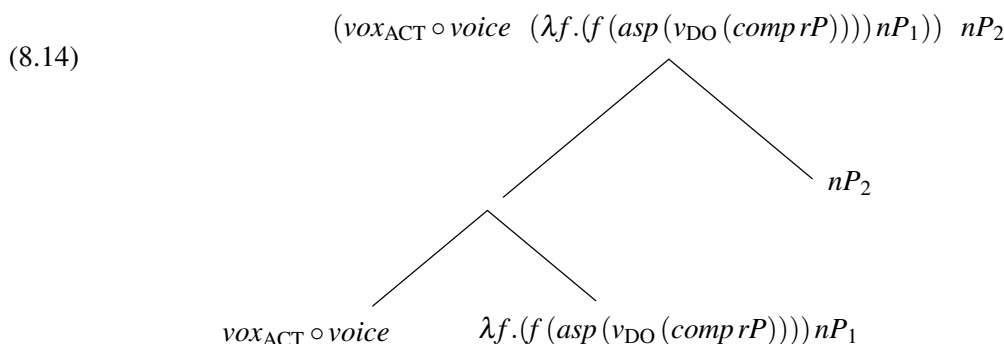
O traço de causalidade *+ec* licencia o argumento nP_2 , introduzido por *Voice*. nP_2 habilita-se como iniciador da eventualidade denotada pelo “constituente inteiro” (não mais apenas o iniciador da eventualidade denotada pelo componente verbal). Ou seja, nP_2 qualifica-se como sujeito ativo do constituinte: o constituinte é compatível com a voz ativa. Deve-se observar que nP_2 é o argumento externo de um constituinte categorizado por V_{DO} , ocupando portanto a posição canônica de um sujeito verbal ativo. Produz-se, então, a derivação mostrada a seguir.⁸

(8.13)



⁸A derivação é representada em dois diagramas, esse e o próximo, por razões práticas, de espaço na página. A visualização do diagrama é importante por apresentar graficamente a história derivacional do constituinte, já codificada em sua própria estrutura.

Pelas considerações acima, sabemos que nP_2 é o sujeito estrutural. Analisemos, agora, a estrutura funcional interna do constituinte derivado, desconsiderando a variável f , para caracterização de nP_1 como objeto estrutural:



O constituinte nominal nP_2 é argumento externo do operador composto $vox_{ACT} voice$, a analisada posição de sujeito estrutural. E “todo” o constituinte $\lambda f.(f(asp(v_{DO}(comprP))))nP_1$ é o argumento interno do mesmo operador. Tem-se, assim, a caracterização do objeto estrutural: o *objeto estrutural* (nP_1) é o “argumento externo” do operador verbalizador ($asp \circ v_{DO}$), se o constituinte verbal é o argumento interno de um núcleo atribuidor de voz. Deve-se observar que a caracterização do objeto estrutural independe dos núcleos situados acima de *Vox* na derivação; tais núcleos contribuem, dependendo da língua, com semântica de tempo, evidencialidade, operações de concordância, atribuição morfossintática de caso, seleção de auxiliares, etc., mas não caracterizam nem o sujeito nem o objeto estruturais.

8.6 Verbos alternantes causativo-incoativos

8.6.1 Alternância de verbos ativos

Até o momento, as análises tiveram em vista as raízes que dão origem a estruturas verbais *ou* transitivas *ou* intransitivas, deixando de lado as raízes que possibilitam a geração de estruturas transitivas *e também* intransitivas, como as dos dois primeiros exemplos a seguir:

- (39)
- a. O rádio quebrou
 - b. João quebrou o radio
 - c. Rádio quebrado
 - d. O rádio está quebrado

Verbos como *quebrar* e *virar* apresentam o fenômeno da “alternância de transitividade” de um modo específico: o *sujeito* da sentença intransitiva é o *objeto* da sentença transitiva. Semanticamente, verbos que participam desta alternância caracterizam-se também como “verbos de

mudança de estado” ou “incoativos”;⁹ e a alternância é conhecida como “alternância causativa-incoativa”.

Tomaremos *quebrar* como representativo desta classe de verbos. Em que pese o fato de *quebrar* originar estruturas interpretáveis como de “mudança de estado”, considerarei o verbo como ativo; isto é, derivado de uma raiz portadora do traço *+ac*. Um forte motivo é o tipo de estatividade apresentada pelo verbo, sempre obtida por meio de um particípio, *quebrado/a*. Neste caso, a estatividade pressupõe sempre uma formação verbal prévia, necessária à geração do particípio. Portanto, a estatividade de *quebrar* é derivada de um processo de verbalização, e não lexicalmente da raiz. A semântica ativa do verbo parece ocorrer também em francês, em que *casser/quebrar* é usado com o auxiliar *avoir/ter*, característico dos verbos “de ação”.

Assim, em verbos como *quebrar* não há uma estatividade proveniente diretamente da raiz, contrariamente ao que se observa de *azedar*: o par *azedar/lazedo* evidencia a estatividade derivada diretamente da raiz (estatividade lexical) coexistindo com a estatividade resultante de um processo de verbalização, *azedar/lazedado*.

Deste modo, toda uma classe de verbos de “mudança de estado”, como *abrir*, *fechar* e *queimar*, será considerada como composta de verbos ativos, derivados de raízes possuidoras do traço semântico *+ac*.

Caso as considerações acima estejam corretas, a incoatividade de um verbo independe da estatividade/atividade da raiz, sendo resultante de um processo mais complexo. Ademais, a designação “alternância causativa-incoativa” é inexata, uma vez que a variante transitiva (causativa) da alternância (‘João quebrou o rádio’) é também incoativa. Por tradição, entretanto, a designação será mantida.

Nesta seção, analisaremos os verbos alternantes ativos e, na próxima, os alternantes estativos.

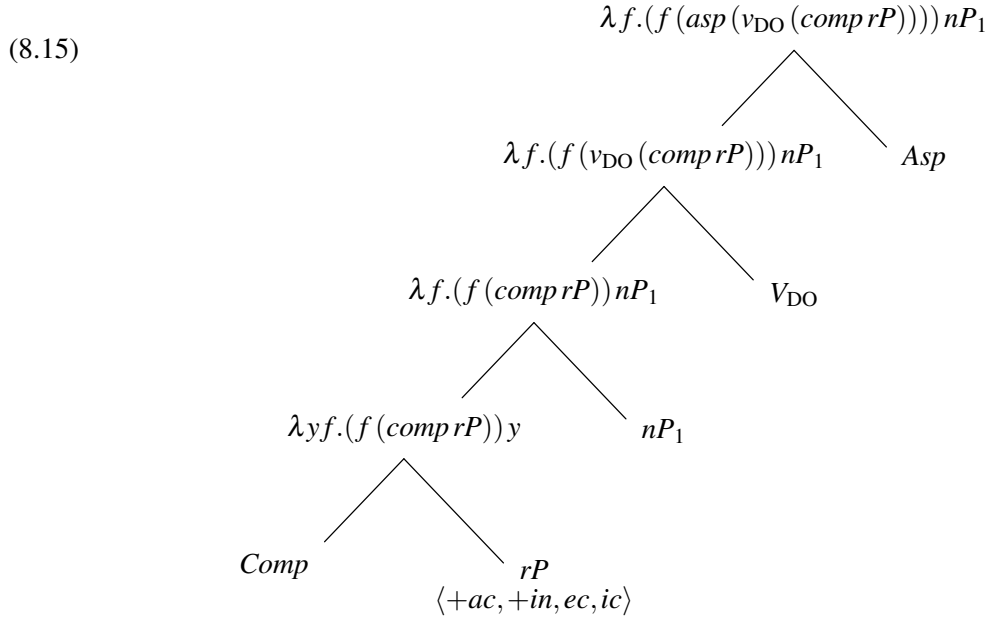
Segundo as análises acima, uma raiz que dá origem a verbos ativos alternantes causativo-incoativos, é possuidora do traço semântico *+ac*. A variante transitiva do verbo (‘João quebrou o rádio’) exhibe compatibilidade da raiz com os traços de incidência, *in* ou *+in*, e de causalidade externa, *ec* ou *+ec*. E a variante intransitiva (‘O rádio quebrou’) mostra que a incidência requerida pela raiz é obrigatória (traço *+in*), e também que existe causalidade interna, fato que, juntamente com a causalidade externa, dota a raiz dos traços *ec* e *ic*. Portanto, as raízes em questão são portadoras dos traços semânticos

$$\langle +ac, +in, ec, ic \rangle.$$

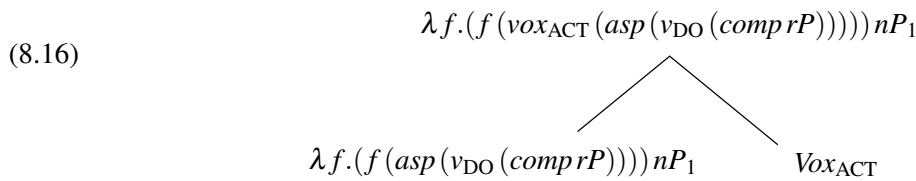
Tal combinação de traços semânticos dá origem à derivação representada no próximo diagrama: (a) o traço semântico *+in* faz a raiz compatível com o núcleo *Comp*, introdutor do argumento *nP₁*, o qual, situado abaixo de *V_{DO}*, torna-se o argumento interno do verbo; (b) qualquer dos

⁹A “mudança de estado” é um dos sentidos do termo *incoativo*; *incoatividade* também tem o significado de “iniciação de um processo”, como em *amanhecer*.

traços de causalidade, *ec* ou *ic*, possibilita a verbalização ativa, por V_{DO} ; e (c) a estrutura verbalizada licencia o núcleo *Asp*.

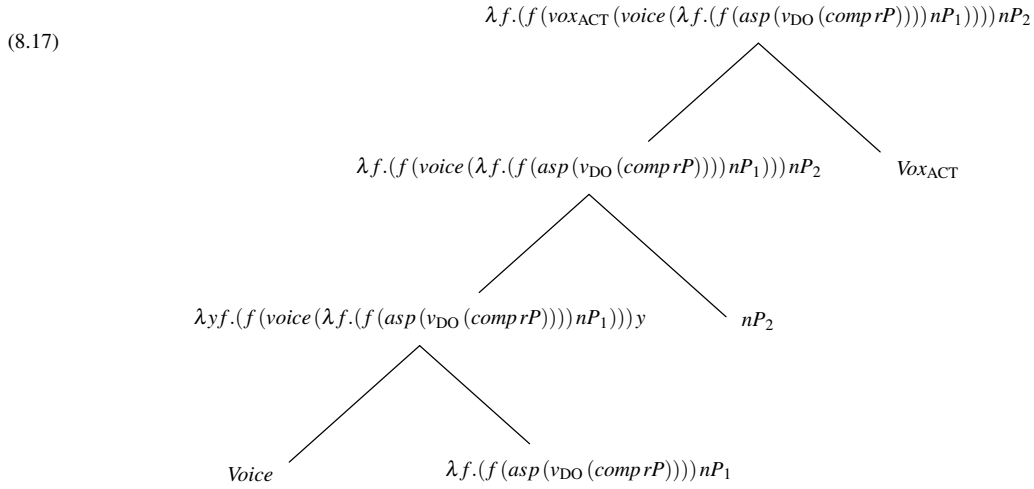


Embora o constituinte $\lambda f.(f(asp(v_{DO}(comp\ rP))))nP_1$ possa ser gerado a partir de *ec* ou *ic*, indistintamente, a continuação da derivação, em cada um dos casos, toma caminhos muito diferentes. O traço de causalidade interna permite a interpretação do argumento interno nP_1 como sujeito estrutural, por meio da voz ativa, de acordo com a derivação mostrada a seguir.



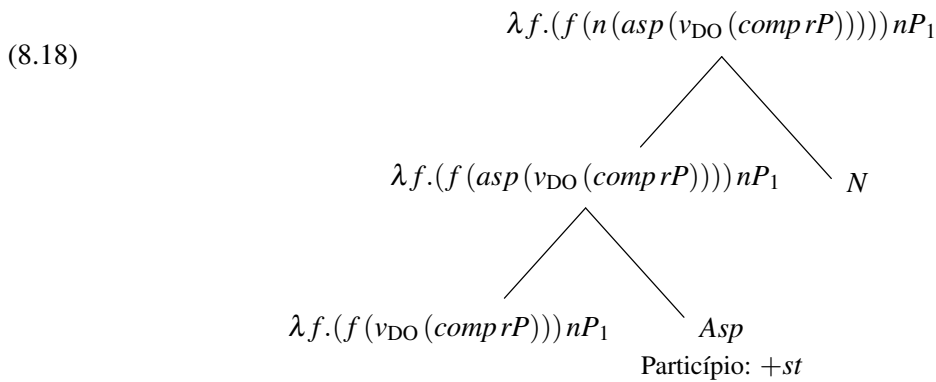
Deriva-se, assim, a variante intransitiva do verbo alternante causativo-incoativo; no caso de *quebrar*, 'O rádio quebrou'. Mais uma vez, é importante observar que nP_1 , mesmo sendo um argumento verbal interno, ocupa a posição estrutural de sujeito, sem que a condição de sujeito tenha resultado de movimento; ao longo de toda a derivação, o argumento permaneceu na posição em que foi inicialmente nela introduzido.

A derivação da variante transitiva resulta do licenciamento do núcleo *Voice* pelo traço *+ec*, o que leva à introdução do argumento nP_2 e à consequente qualificação deste argumento como sujeito estrutural pela voz ativa, produzindo-se a derivação:

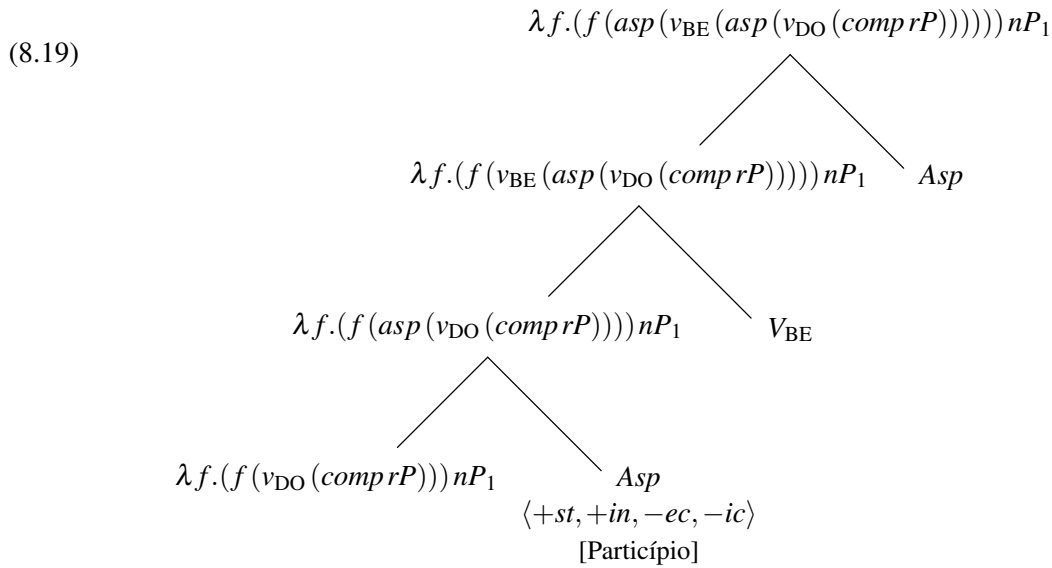


Tem-se, assim, a derivação de ‘João quebrou o rádio’, a variante transitiva do verbo alternante. O sujeito, nP_2 /João, é o argumento externo do verbo, introduzido por *Voice*, e o objeto estrutural, nP_1 /rádio, é o argumento verbal interno, introduzido por *Comp*.

O constituinte $\lambda f.(f(asp(v_{DO}(comp rP))))n P_1$, que gerou a variante verbal intransitiva, por meio de vox_{ACT} , não está comprometido com uma interpretação verbal, podendo alternativamente derivar constituintes nominais, como ‘Rádio quebrado’, ao ser nominalizado pelo núcleo *N*, de acordo com a derivação mostrada no diagrama a seguir.



A derivação de ‘O rádio está quebrado’ requer a formação prévia do particípio. De acordo com a proposta, o particípio é um aspecto verbal estativo, destituído de causalidade, fato que cria as condições necessárias à verbalização estativa por V_{BE} . Como se observa no constituinte situado no topo do diagrama abaixo, o operador v_{BE} “enxerga” em *asp*, primeiro operador do seu argumento interno, a configuração de traços $\langle +st, in, -ec, -ic \rangle$ característica de *Asp* ao assumir o valor de particípio.



Nas derivações acima, a alternância causativo-incoativa resulta da co-ocorrência dos traços *ec* e *ic* em colaboração com os núcleos *Vox* e *Voice*; isto é, resulta da informação lexical codificada na raiz, em combinação com núcleos (por definição, gramaticais). A raiz é agnóstica com relação às estruturas em que pode ser inserida; ela se apresenta aos núcleos como um repositório de informações, basicamente os traços semânticos e secundariamente os traços fonológicos; as estruturas possíveis decorrem da compatibilidade dos núcleos com os traços dos constituintes a que eles se aplicam.

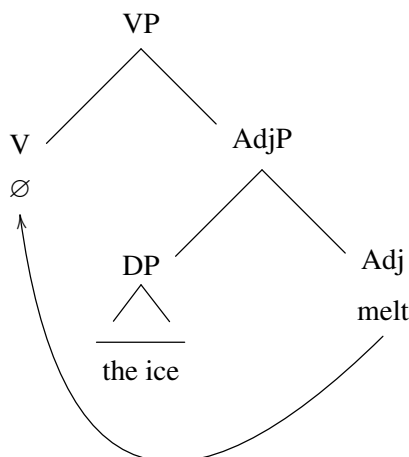
Portanto, as alternâncias de transitividade (como todos os demais fenômenos sintáticos) derivam da colaboração entre núcleos e traços semânticos; as configurações sintáticas são obtidas por consequência, não sendo diretamente especificadas nas raízes nem em qualquer outro constituinte. A presente proposta difere assim radicalmente da de Hale e Keyser, para quem a alternância de transitividade se deve a propriedades lexicais das raízes, mas em um sentido diferente do suposto neste trabalho. Para os dois autores, as raízes codificam diretamente a informação a respeito das configurações sintáticas em que podem ser inseridas.

Verbos ativos alternantes e estatividade

A interpretação dos verbos inacusativos de mudança de estado como verbos necessariamente estativos leva a incongruências entre a representação desses verbos e a de outras estruturas a eles relacionadas.

Ao procurar unificar a representação sintática dos verbos inacusativos de mudança de estado, Harley, reelaborando ideias de Hale e Keyser, propõe que a estrutura verbal, em especial a da fonologia, é formada por *conflation* entre um núcleo verbal e um adjetivo, segundo a estrutura mostrada no próximo diagrama (Harley 2002a, p. 8).

(8.20)



A solução de Harley tem em vista verbos como *clear/clarear* e *melt/derreter*, que dão origem às estruturas a seguir.

- | | | |
|-----|----|--|
| (1) | a. | Mary cleared the screen/Maria clareou a tela |
| | b. | The screen cleared/A tela clareou |
| (2) | c. | Mary melted the ice/Maria derreteu o gelo |
| | d. | The ice melted/O gelo derreteu |
| (3) | e. | The screen is clear/A tela está clara |
| | f. | The ice is melted/O gelo está derretido |

Nestes exemplos, além das alternâncias de transitividade, o fenômeno mais significativo é a diferença de fonologia dos adjetivos, derivada diretamente da raiz verbal, no caso de *clear*, e derivada de um verbo já formado, no caso de *melted*. Considerado o diagrama, tal fato evidencia uma dificuldade na proposta de Harley: enquanto a fonologia do verbo resulta bem-formada no caso de *clear*, há uma clara inadequação no caso de *melted*, em que o sufixo verbal *ed* precisa ser desconsiderado. Harley (2002a, p. 8) atribui o problema a uma inexplicada “lacuna morfológica”: [a representação do diagrama acima] “... makes melt the ice consistent with other unaccusative change-of-state verbs, at the cost of assuming a gap in the morphology: there’s no underived adjectival form of melt.”¹⁰

Ao abordar as alternâncias de transitividade como as dos exemplos anteriores, Hale e Keyser, em uma proposta pioneira, também baseiam suas soluções no mecanismo de *conflation* entre o núcleo verbal e um adjetivo. Entretanto os dois pesquisadores ressaltam, em uma das análises, que a solução apresentada para determinado caso básico (a variante transitiva de *clear*) efetivamente não se verifica em representações sintáticas, uma afirmação sem dúvida desconcertante: “... we should mention that representations like [those] do not exist at any stage in the

¹⁰“...faz melt the ice [derreter o gelo] consistente com outros verbos inacusativos de mudança-de-estado, à custa de assumir uma lacuna na morfologia: não existe forma adjetival não-derivada para melt [derreter]”.

syntactic representations of sentences.”, Hale & Keyser (2002, p. 17)¹¹

Assim, tanto nas soluções de Harley quanto nas de H&K, existem inconsistências entre as estruturas apresentadas para representar os verbos incoativos alternantes e as estruturas propostas para representar as demais configurações verbais.

Grande parte das inconsistências provém da interpretação dos verbos de “mudança de estado” como necessariamente estativos. Se agregarmos o fato de, em função da suposta estatividade, esses verbos serem considerados deadjetivais (derivados de adjetivos), encontraremos dois pressupostos que, em algum momento, não deixarão de acarretar má-formação da estrutura verbal: as dificuldades não são acidentais.

De acordo com a presente pesquisa, como vimos nas páginas anteriores, as alternâncias de transitividade e a incoatividade não constituem fenômenos originalmente verbais, resultam antes dos traços de causalidade da raiz, sendo ademais independentes dos traços de estatividade.

A solução que proponho para *melt/melted* segue o modelo geral dos verbos ativos alternantes, como *quebrar/quebrado*, cujo padrão estrutural se encontra analisado na seção 8.6.1, p. 112. Neste tipo de verbos, toda a família de estruturas relacionadas (incoativa, transitiva, nominal e verbal estativa) apresenta derivações inteiramente regulares e compatíveis com as das demais estruturas verbais, sem exibir lacunas morfológicas.

8.6.2 Alternância de verbos estativos

Analisaremos agora a alternância causativo-incoativa de verbos como *esfriar*, *azedar*, *avermelhar*, *esverdear*, *escurecer* e *engordar*, derivados de raízes estativas. O verbo *esfriar* será tomado como representante desta classe e a alternância é mostrada nos dois primeiros exemplos a seguir.

- | | | |
|-----|----|----------------------|
| (4) | a. | O café esfriou |
| | b. | Maria esfriou o café |
| | c. | Café frio |
| | d. | Café esfriado |
| | e. | O café está frio |
| | f. | O café está esfriado |

De acordo com as análises da seção 7.1, p. 88 e p. 89, as raízes dos verbos acima são portadoras dos traços semânticos $\langle +st, in, -ec, -ic \rangle$, sendo, assim, raízes de estatividade básica. A alternância de transitividade coloca a questão: como os traços $-ec$ e $-ic$, incompatíveis com o verbalizador ativo, V_{DO} , podem dar origem a uma sentença de atividade, como ‘Maria esfriou o café’, e a outra em que há manifestação de mudança, como ‘O café esfriou’? De que modo raízes

¹¹“...devemos mencionar que representações como [aquelas] não existem em nenhum estágio das representações sintáticas de sentenças.”

puramente estativas, somente compatíveis com V_{BE} , derivam sentenças em que há elementos de semântica ativa?

A resposta é dada pela morfologia do verbo *esfriar*, em que o morfema *es* evidencia a existência de um processo atuante sobre a raiz. Na quase totalidade dos verbos citados como exemplos (*avermelhar*, *esverdear*, *escurecer*, *engordar*), há mostras de alguma morfologia-extra aplicada à raiz.

Existem aparentes exceções, a serem analisadas algumas páginas adiante, como *azedar* e *alegrar*, e casos bastante desafiadores, como as raízes de *alto* e *bom*, que não dão origem diretamente a verbos.

Comparemos com *chutar*, *cair*, *chegar* e *quebrar*, em que os acréscimos fonológicos à raiz são resultantes de processos claramente pertencentes ao sistema verbal: (a) a vogal temática, própria do processo de verbalização; e (b) a expressão do infinitivo. Tais verbos exibem estatividade derivada (*chutado*, *caído*, *chegado*, *quebrado*); isto é: (a) estatividade não lexicalmente codificada diretamente na raiz, (b) não se tratando, assim, de raízes de estatividade pura.

Descontemos esses dois tipos de acréscimo (vogal temática e marca de infinitivo) aos verbos da classe de *esfriar*. Podemos, então, supor que os acréscimos fonológicos (quase todos prefixados) às raízes de estatividade pura, associados à verbalização, sinalizam a existência de um processo não estritamente verbal aplicado à raiz, mas necessário à verbalização.

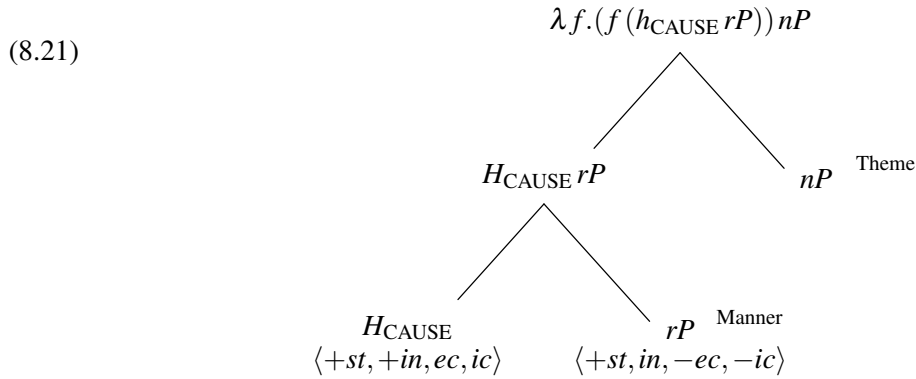
Nos termos da presente proposta, um tal processo é implementado por um núcleo aplicado à raiz. A questão é então determinar as características semânticas e fonológicas deste núcleo, uma vez que o seu comportamento sintático será regido pelos princípios gerais do modelo computacional. Podemos determinar a semântica do núcleo a partir das estruturas a que ele dá origem.

O núcleo:

- (a) Estabelece a associação entre uma raiz de traços $\langle +st, in, -ec, -ic \rangle$ e um *nP*.
- (b) Possui semântica estativa: a semântica das construções intransitivas 'O café esfriou', 'A sala escureceu', etc., é claramente estativa.
- (c) Requer incidência necessária: a raiz deve obrigatoriamente ligar-se a um *nP*.
- (d) Contém os traços *ec* e *ic*, necessários à geração das estruturas transitiva e intransitiva.

Este núcleo será denotado por H_{CAUSE} ,¹² uma vez que sua atuação fundamental consiste em acrescentar traços de causalidade a uma raiz de estatividade básica. Gráficamente, o núcleo possui a configuração:

¹²A designação *H* foi adotada para reservar o uso de *P* exclusivamente a preposições, núcleos que estabelecem uma associação entre dois *nPs*.



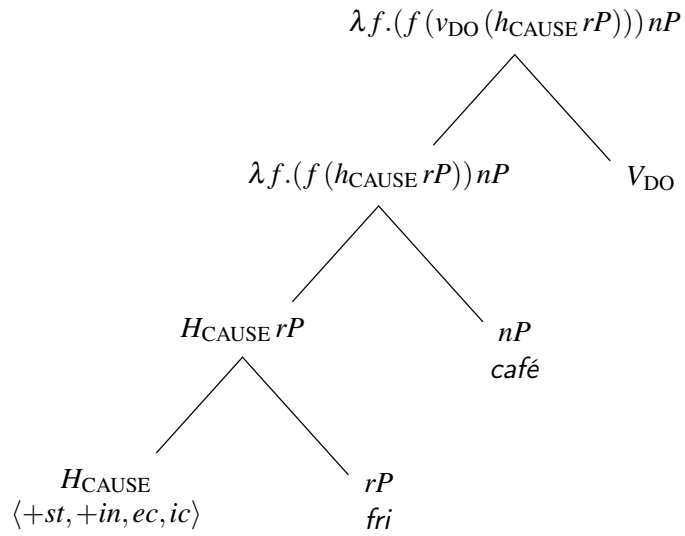
Por propriedades gerais dos núcleos, os traços *ec* ou *ic* de H_{CAUSE} licenciam a verbalização ativa, por V_{DO} . Sob a ação do núcleo, a raiz é posta em uma condição de incidência sobre o nP ; assim, H_{CAUSE} introduz uma interpretação de rP como *modo* e de nP como *tema* (não necessariamente afetado), fato formalizado por meio dos papéis semânticos Manner e Theme. Deve-se ressaltar que os papéis semânticos são interpretações de posições estruturais, não se encontrando registrados nem no núcleo nem em nenhum de seus argumentos.¹³ No modelo que proponho, os papéis semânticos (temáticos) são estruturalmente determinados, não fazendo parte do conteúdo lexical de um constituinte.

A variante intransitiva de *esfriar*, ‘O café esfriou’, tem a sua derivação representada no próximo diagrama, em que foram omitidos os núcleos introduzidos na derivação acima de V_{DO} , como *Asp* e *Vox*. O constituinte situado no topo do diagrama é composto pela cadeia de operadores $f \circ v_{\text{DO}} \circ h_{\text{CAUSE}}$, juntamente com seus dois argumentos, rP e nP . Deve-se ter em conta que f contém a cadeia dos núcleos omitidos da derivação. Sob tais condições, o nP , *argumento externo* da cadeia de operadores, é licenciado como sujeito da estrutura pelo traço de causalidade interna, *ic*, de H_{CAUSE} . É importante observar que o nP recebe o tratamento de sujeito sem mover-se da posição em que foi introduzido na estrutura.

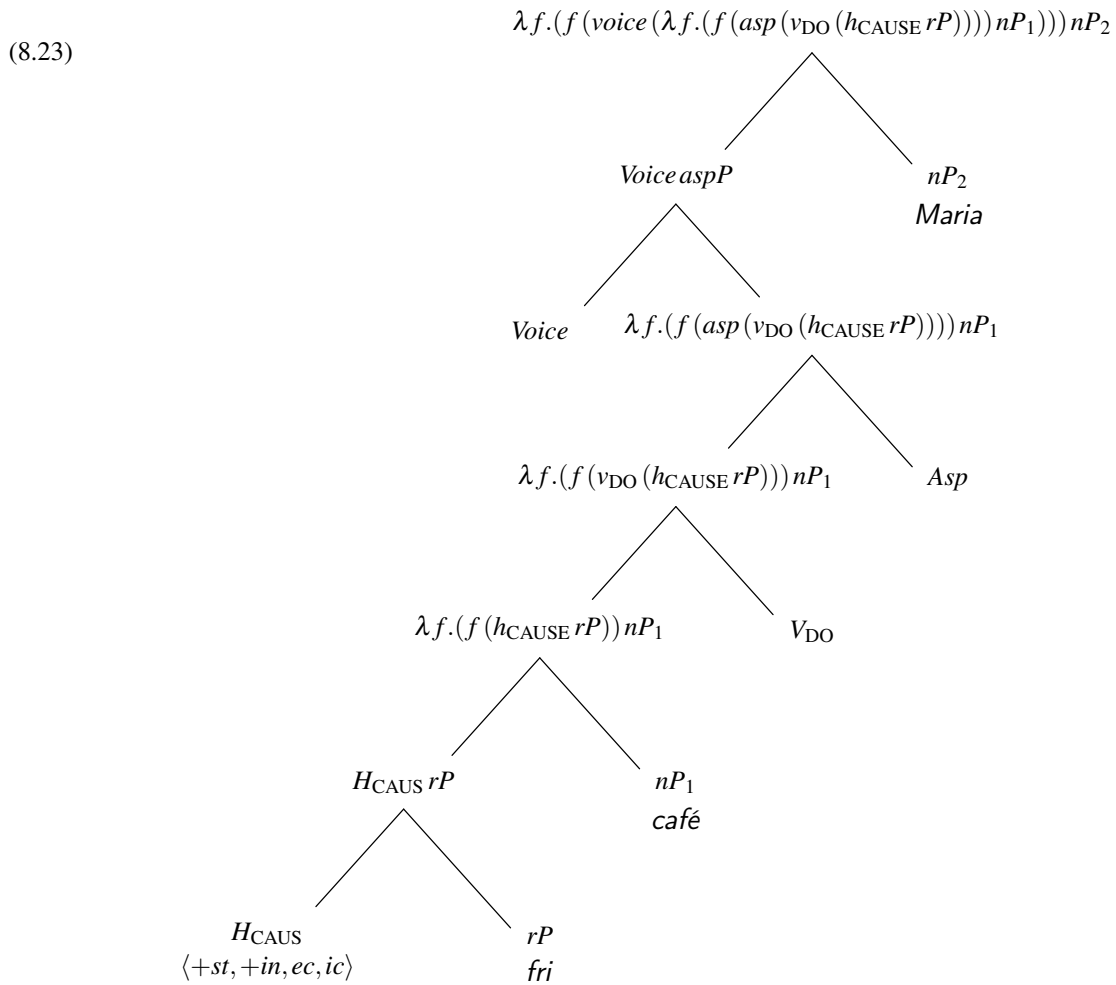
A estrutura causativa-intransitiva mostra um dos modos como um *argumento interno* ao processo de verbalização (um argumento introduzido na derivação abaixo do núcleo verbalizador, no caso, V_{DO}) recebe o tratamento estrutural de sujeito.

¹³ A expressão “papel temático” fica reservada para o seu uso clássico na teoria gerativa, o de registro específico em uma entrada lexical.

(8.22)



Além da compatibilidade com sujeitos internos, o traço *ec*, de causalidade externa, torna H_{CAUSE} compatível com o núcleo *Voice*, introdutor de argumentos externos. O próximo diagrama mostra a derivação de ‘Maria esfriou o café’, representante da estrutura causativa-transitiva. Nesta estrutura, a cadeia de operadores $f \circ \text{voice}$ tem: (a) o constituinte nP_2 como argumento externo, situado na posição estrutural de sujeito; e (b) o constituinte verbal, encabeçado por $f \circ v_{\text{DO}}$, como argumento interno, o que confere a nP_1 a condição de objeto estrutural, independente do traço *ic*.



Semântica de H_{CAUSE}

O núcleo H_{CAUSE} aplica-se a raízes isentas de causalidade, portadoras dos traços

$$\langle +st, in, -ec, -ic \rangle.$$

Estados não-causados (dotados dos traços $-ec$ e $-ic$) independem de um ato de iniciação, situando-se, de certo modo, fora do tempo. Assim, as raízes estativas básicas são especialmente aptas a denotar propriedades inerentes ('folha verde', 'folha vermelha') ou estáticas ('café frio', 'café quente'). Este é o domínio natural do verbalizador V_{BE} .

Por outro lado, os traços de causalidade (contingente: ec e ic ; necessária: $+ec$ e $+ic$) caracterizam ocorrências, pondo em cena a ideia de *início* de algo, ou de *antes* e *depois*. As raízes possuidoras desses traços individualizam processos, situações dotadas de algum dinamismo. O núcleo V_{DO} é o formalizador dos eventos em que está implícito o surgimento ou a mudança.

O núcleo H_{CAUSE} acrescenta o elemento de causalidade a uma raiz que denota permanência estática, instituindo a noção de início de uma situação. Esta é a semântica básica da incoatividade. H_{CAUSE} introduz a ideia de iniciação de um estado, de ocorrência de uma mudança. Ou seja, H_{CAUSE} é responsável pela interpretação incoativa da raiz. A verbalização por V_{DO} forma então um “verbo de mudança de estado”.

Deve-se observar que a semântica de mudança é dada por H_{CAUSE} , e não por V_{DO} , um núcleo verbalizador. Não há algo semelhante a ‘X BECOME Adj’, como em Harley (2002a, 7), em que o verbalizador se expressa gramaticalmente por BECOME ao ter um adjetivo como complemento. Cada núcleo proposto introduz um único efeito gramatical; não há como um verbalizador, como V_{DO} , acrescentar uma semântica-extra, como a incoatividade, o que deve ficar a cargo de um núcleo específico, no caso H_{CAUSE} ; ademais, não há uma categoria lexical primitiva de adjetivo em minha proposta.

Expressão fonológica de H_{CAUSE}

Todo núcleo possui expressão fonológica puramente gramatical, determinada pelo seu paradigma interno, segundo o contexto em que está incluído. O morfema *es*, de *esfriar*, é um prefixo, especificado pela lógica interna de H_{CAUSE} , e anexado à fonologia da raiz *fri*; é o mesmo prefixo presente em *esquentar*.

Normalmente um núcleo se expressa fonologicamente de modo variado. Em português, V_{BE} apresenta uma riquíssima morfossintaxe, dependendo do aspecto, tempo, modo e pessoa. A preposição que implementa a relação genitiva em inglês (*of*) se expressa fonologicamente como preposição, como desinência, ou até mesmo por meio da ordem dos constituintes. Fatores de contexto sintático, semântico, lexical, estilístico, e também histórico, condicionam a expressão fonológica de um núcleo.

O núcleo H_{CAUSE} , em português, não é exceção; pode expressar-se de diferentes modos, na maioria das vezes por prefixação. Por exemplo: (a) prefixação de *a* – *avermelhar*, *arroxear*; (b) prefixação de *es* – *esquentar*, *esfriar*, *esverdear*; (c) prefixação de *em* – *engordar*, *emagrecer*; (d) prefixação e sufixação – *escurecer*, *entristecer*.¹⁴

O objetivo dos exemplos acima, evidentemente, não é fazer um estudo morfológico de uma classe de verbos incoativos em português, mas ilustrar a expressão fonológica de um núcleo proposto, e, principalmente, obter material comparativo para a análise de raízes em que o núcleo não se aplica, ou parece não se aplicar.

Uma observação é necessária: foram utilizados apenas verbos para ilustrar a expressão de H_{CAUSE} , entretanto, o núcleo aplica-se a raízes. O motivo é o núcleo somente tornar-se visível mediante a verbalização por V_{DO} .

Aparentemente H_{CAUSE} não se aplica aos verbos *amarelar* e *azedar*: os verbos par-

¹⁴Há uma interpretação da morfologia de *esfriar* (e analogamente da dos demais verbos exemplificados acima) em termos de uma operação de circunfixação (*es* ... *ar*). Uma tal solução, em termos puramente morfológicos, deixa algumas questões em aberto: por que o *ar* de *esfriar* faria parte de um circunfixo e o *ar* de *falar* não o faria?

ticipam das alternâncias de transitividade ('O repolho azedou'/'Maria azedou o repolho'), mas as raízes permanecem inalteradas, diferentemente do que ocorre a *avermelhar*, em que a fonologia da raiz recebe um prefixo. Entretanto, *amarelar* e *azedar* exibem exatamente o mesmo comportamento sintático de *avermelhar*, não apenas nas alternâncias de transitividade, mas também na voz média (um tema ainda a ser desenvolvido). A não-alteração das raízes *azed* e *amarel* é um fenômeno puramente fonológico, devido à vogal inicial de ambas as raízes.

Situação diferente tem-se com as raízes de *alto*, *bom* e *feliz*: (a) são raízes de estatividade básica, sem (b) qualquer irregularidade quanto às construções estativas ('Maria é alta/boa/feliz'); (c) inexistem, porém, as construções incoativas *'Maria alteia/boneia/feliza'.

Mas é possível levar as análises um pouco mais adiante. Pode-se parafrasear 'O café esfriou' por 'O café tornou-se mais frio';¹⁵

do mesmo modo, e, além disso, 'Maria tornou-se alta/boa/feliz' é gramatical. Ademais, também existe a modalidade incoativa-transitiva 'João tornou Maria alta/boa/feliz'.¹⁶ Tal possibilidade revela um fato: o núcleo H_{CAUSE} possui uma expressão fonológica por *default*, 'tornar[-se] mais ...', sempre aplicável às raízes estativas básicas. Ocorre que, por razões lexicais não contempladas nos traços *st*, *in*, *ec* e *ic*, algumas raízes não são compatíveis com as fonologias de H_{CAUSE} "morfologicamente" manifestadas por meio de afixos. Do mesmo modo que há raízes lexicalmente compatíveis com os prefixos *a*, *es*, *em*, etc., também há as raízes lexicalmente incompatíveis com a morfologia do núcleo expressa por meio de afixos. Portanto, o núcleo causativo H_{CAUSE} , caso as análises estejam corretas, é universalmente aplicável às raízes de estatividade básica.

Deve-se observar que H_{CAUSE} não pode ser confundido com o causativo sintático costumeiramente manifestado por *fazer*: 'João fez Maria feliz/rir/chegar/fritar o ovo'. Expressões como 'João fez Maria tornar-se feliz' e a agramaticalidade de 'João tornou Maria fazer[-se] feliz' mostram que H_{CAUSE} e o causativo sintático não são o mesmo núcleo. Em uma derivação, o causativo sintático expresso por *fazer* ocorre acima de *Comp* e do núcleo verbalizador, enquanto que H_{CAUSE} aplica-se diretamente a raízes.

A afirmação de que H_{CAUSE} se aplica universalmente às raízes de estatividade básica exige que algumas estruturas sejam analisadas com cuidado. O que se pode dizer das sentenças abaixo?

- (5) a. O barco afundou/ O capitão afundou o barco

Tem-se uma sentença incoativa alternante, e, no entanto, 'O barco afundou' não pode ser parafraseada por 'O barco tornou-se mais fundo'. O exemplo parece contradizer a própria (suposta) realidade do núcleo H_{CAUSE} .

Entretanto, devemos considerar que 'O barco afundou' é paráfrase de 'O barco foi para

¹⁵ Agradecimentos ao Luiz Arthur Pagani, por observar que a paráfrase não poderia ser 'O café esfriou', como originalmente proposto: "o café pode ter-se esfriado sem ter-se tornado frio". A incoatividade, efetivamente, não pressupõe um estado final realizado.

¹⁶ 'João tornou Maria alta, ao dar-lhe um sapato de salto.'

o fundo'. Ou seja, *afundar* não possui a mesma semântica de *avermelhar* ou de *engordar*. Enquanto em *avermelhar* um *nP* encontra-se em relação imediata com uma raiz, em *afundar* estruturalmente tem-se a relação entre dois *nPs* (*barco* e *fundo*). A estrutura de *afundar*, *engavetar*, etc., deriva de preposições cujos argumentos internos denotam locais, analisadas em Hale & Keyser (2002), e que serão assunto da seção 8.8.7, p. 144. De todo modo, 'O barco tornou-se mais fundo', uma das expressões de H_{CAUSE} , difere de 'O barco foi para o fundo', expressão estrutural possível da preposição. Não se tem, assim, a mesma estrutura em *afundar* e *avermelhar*, embora a expressão fonológica dos núcleos envolvidos seja a mesma.¹⁷

Muitas vezes, a analogia das formas pronunciadas pode levar a soluções inadequadas quando se analisam estruturas de constituintes. Um bom princípio orientador é sempre analisar famílias de estruturas relacionadas, sobretudo buscando exemplos que possam invalidar as soluções propostas.

8.6.3 Alternância, transitividade e incoatividade

Há, portanto, duas classes de verbos que exibem a alternância causativa-incoativa: verbos como *quebrar*, em que a alternância se deve a propriedades lexicais da raiz; e como *esfriar*, cuja alternância resulta de um núcleo causativo aplicado à raiz. Em ambos os casos, porém, a verbalização por V_{DO} , necessária à alternância, ocorre sobre estruturas similares, um núcleo que se aplica a um *rP* e um *nP*.

Já a derivação das estruturas estativas, por V_{BE} , difere radicalmente em cada caso. Em *esfriar*, a verbalização estativa se aplica acima de uma raiz não-categorizada, enquanto que em *quebrar* a mesma verbalização somente é possível acima de V_{DO} , o que torna o particípio inevitável.

No meu modelo, a alternância de transitividade depende apenas da conjunção dos traços *ec* e *ic*, não de *st* ou *ac*. Isto implica que tanto verbos ativos quanto estativos podem exibir comportamento alternante.

Pode ser que o caráter alegadamente estativo dos verbos do tipo de *quebrar* (incoativos alternantes lexicais) se deva à possibilidade de formação da construção intransitiva, em que o sujeito é "afetado pelo verbo" e é não-agentivo. Ou seja, a estatividade do verbo estaria genericamente associada ao caráter inacusativo da construção intransitiva, os inacusativos sendo interpretados como estativos.

Os verbos alternantes lexicais (ativos) seriam estativos por apresentarem uma versão intransitiva; isto diferenciaria *quebrar* de *cortar*. Ao menos no modelo que proponho, não há diferença entre a derivação transitiva de um verbo alternante e a derivação de um verbo transitivo; a única diferença reside na semântica da raiz, *—ic* no primeiro caso, e *ic* no segundo. A diferença reside, assim, na possibilidade de o argumento interno ser interpretado como sujeito.

¹⁷Deste modo, *avermelhar* e *afundar* não têm "a mesma morfologia": as expressões fonológicas dos núcleos são análogas, mas as estruturas subjacentes são diferentes.

Dito de outro modo: os verbos alternantes seriam estativos por serem alternantes; a estatividade teria sua origem em um fenômeno interpretativo, resultante da condição peculiar do sujeito da variante intransitiva.

O caráter incoativo (mudança de estado) dos verbos alternantes se deve à verbalização ativa, V_{DO} , que introduz uma noção de atividade associada à condição do sujeito intransitivo, sendo, portanto, um fenômeno interpretativo, um efeito colateral da condição do sujeito.

No modelo que proponho, não há ligação necessária entre alternância de transitividade e estatividade, nem entre inacusatividade e estatividade.

Uma raiz claramente estativa como *fri* gera uma construção estativa “pura”, como ‘café frio’ e a outra derivada, como ‘café esfriado’, em que há traços de atividade. Mas a raiz *quebr* dá origem apenas a ‘rádio quebrado’, do mesmo modo que a raiz *chut*, inquestionavelmente ativa, somente origina expressões como ‘bola chutada’. Há, assim, muitos argumentos favoráveis a se considerarem as raízes de muitos verbos incoativos alternantes como ativas. Neste caso, a incoatividade de tais verbos seria um fenômeno interpretativo, ligado à própria alternância de transitividade, e não à semântica estativa ou ativa da raiz.

De todo modo, quero frisar que, no meu modelo, a alternância de transitividade independe da estatividade da raiz, e que verbos podem ser “incoativos alternantes” e, ainda assim, serem derivados de raízes ativas.

8.7 Verbos ditransitivos

Nos capítulos anteriores, enfocamos os primeiros elementos de formação da estrutura argumental, analisando a verbalização estativa básica, os verbos lexicais (aqueles cuja derivação depende diretamente do conteúdo semântico e fonológico da raiz, a saber, inacusativos, inergativos e transitivos “puros”), os verbos incoativos e a alternância causativo-incoativa.

Neste capítulo estenderemos a análise a três outras importantes classes de verbos:

- (a) Os verbos de *duplo-complemento* e de *duplo-objeto*, como, respectivamente, ‘John gave a book to Mary’ e ‘John gave Mary a book’; daremos, genericamente, a tais verbos a denominação de *ditransitivos*;
- (b) Os verbos gramaticais *dar* e *ter*;
- (c) Verbos denominais, como *engavetar* e *enlamear*.

As três classes de verbos possuem propriedades estruturais em comum e, consideradas conjuntamente, constituem um dos temas que mais têm contribuído para o desenvolvimento da teoria sintática.

No âmbito gerativo, a primeira abordagem unificada de representação dos verbos ditransitivos se deve a Chomsky (1955/1975). A partir desta análise inicial, a investigação se alargou para abranger uma ampla gama de fenômenos, proporcionando um tratamento unificado a outras classes de verbos, como os gramaticais e os denominais, mencionados acima. Toda a teoria da gramática gerativa — sintaxe, semântica e morfologia —, em suas várias vertentes, tem sido influenciada pelos conceitos e métodos desenvolvidos para solucionar as questões levantadas por essas estruturas.

Uma das linhas de pesquisa atuais, na qual nosso trabalho se insere, propõe que as três classes de verbos resultam da verbalização de preposições da Gramática Universal.

Neste capítulo faremos uma breve revisão das principais abordagens que contribuíram para a nossa proposta. Começaremos com a apresentação da sugestão de Chomsky para o tratamento transformacional da derivação dos verbos ditransitivos. Em seguida, analisaremos a proposta pioneira de Larson, que, desenvolvendo a proposta de Chomsky, lançou as bases da moderna teoria gerativa da estrutura verbal. Prosseguiremos com o delineamento das soluções de Hale e Keyser, que desenvolvem a projeção sintática dos argumentos e a interpretação temática a partir da estrutura dos constituintes. Passaremos então às propostas de Heidi Harley, que, até onde vejo, constituem a abordagem mais consistente e empiricamente mais bem fundamentada dos fenômenos abordados no capítulo. Analisaremos, para terminar, as soluções da minha pesquisa.

8.7.1 Os verbos biargumentais

Desde as fases iniciais da Gramática Gerativa, a derivação das sentenças com duplo-complemento, como a primeira abaixo, parecia adequadamente resolvida: o argumento preposicionado *him*, representado como um constituinte PP anexado à derivação como adjunto mostrava-se uma solução bastante clara. As transformações, sobretudo a passiva, resultavam do modo esperado, comprovando que a estrutura da sentença se encontrava bem compreendida.

- (6) a. The teacher gave several books to him
- b. The teacher gave him several books
- c. They elected him an officer

Por outro lado, as sentenças com duplo-complemento, como a segunda acima, apresentavam resultados inesperados ao serem submetidas a transformações. Chomsky (1955/1975, p. 493) observa que esta sentença pode ter duas passivas:

- (7) a. He was given several books by the teacher
- b. Several books were given him by the teacher

No entanto, 'They elected him an officer', a terceira sentença acima, somente tem uma passiva

- (8) a. He was elected an officer by them
- b. *An officer was elected him by them

8.7.2 A sugestão de Chomsky

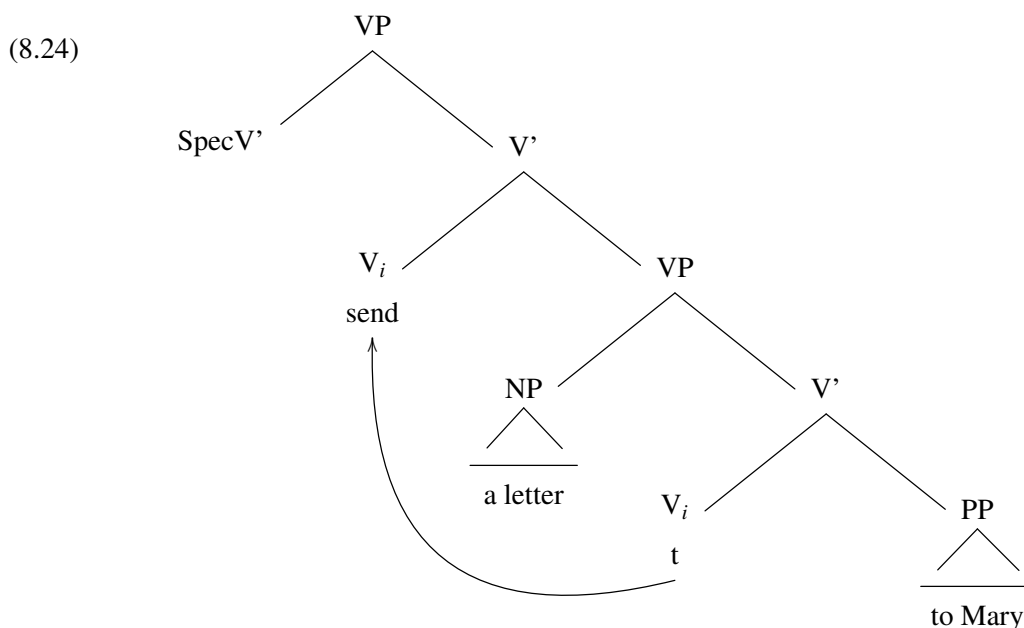
Fenômenos transformacionais deste tipo levaram Chomsky (1955/1975, pp. 492–94), a propor que deveria haver uma sentença mais elementar do que 'The teacher gave him several books': a sentença mais elementar deveria ser 'The teacher gave several books to him', da qual a primeira se derivaria por meio de uma transformação. Na segunda sentença, segundo Chomsky, *to him* não é verdadeiramente um PP, mas uma espécie de objeto interno do verbo, formando com ele um constituinte (*give-him*); este constituinte composto entra em relação de complementação com *several books*, o objeto direto superficial. Deriva-se, assim, 'The teacher gave him several books'.

A solução de Chomsky introduz uma importante ideia para o tratamento de expressões em que figura o dativo com *to* (*to-dative*): o dativo deixa de ser representado por meio de um adjunto, passando a ser interpretado como complemento estrutural do verbo. O desenvolvimento desta ideia mostrar-se-á fundamental para a moderna teoria da sintaxe gerativa.

O próximo passo será dado por Larson.

8.7.3 A proposta de Larson

Larson (1988) desenvolve a abordagem de Chomsky, em que a estrutura de duplo-objeto é derivada da de duplo-complemento, de acordo com a estrutura do diagrama abaixo (Larson 1988, p. 343).



Larson propôs a ideia de que a estrutura argumental de um verbo ditransitivo é projetada por uma cadeia de (dois) núcleos, e que a posição final do verbo, na sentença pronunciada, é obtida por movimento núcleo-a-núcleo. A estrutura resultante, em níveis, tornou-se conhecida como a “*Larsonian shell*”/a concha larsoniana.

Com sua proposta, Larson desenvolve a ideia de Chomsky segundo o paradigma da teoria X-barra, representando a estrutura verbal de um modo estritamente binário (uma restrição nem sempre seguida em muitos desenvolvimentos posteriores).

Larson deu origem também à noção de que o núcleo verbal superior possui função puramente estrutural, necessária à introdução do argumento externo. Assim o fazendo, abriu caminho para que a interpretação semântica (temática) do argumento externo do verbo se desvesse à sua posição estrutural, não se ligando, necessariamente, a uma especificação temática na entrada lexical do verbo.

Entretanto, como observa Harley (2007, p. 5), a proposta de Larson apresenta dificuldades com relação à teoria temática, principalmente quanto à atribuição de papel temático ao argumento externo.

Segundo a presente pesquisa, é como se a introdução estrutural do argumento externo não se harmonizasse adequadamente com prescrições temáticas lexicalmente especificadas. Contudo, creio não haver incompatibilidade fundamental entre a *representação estrutural* de

Larson e os pressupostos lexicalistas de suas análises. A estrutura sintática em camadas pode ser gerada a partir da entrada lexical suposta por Larson, tanto quanto pode sê-lo a partir de outras estruturas de entrada lexical. Se houver consequências problemáticas decorrentes do estilo de representação estrutural de Larson, a causa não será a solução sintática, utilizada ainda nos dias de hoje, nem seguramente será a incompatibilidade da estrutura sintática com a especificação temática lexical. Chomsky, por exemplo, incorporou a *Larsonian shell* à Teoria de Princípios e Parâmetros (TPP) (Chomsky 1995, p. 62) e a manteve no Programa Minimalista (PM) (Chomsky 1995, p. 180, 307). Qualquer questionamento à solução de Larson seguramente terá por objeto a própria teoria temática por ele adotada, a teoria de Chomsky (1981) reelaborada por Marantz (1981):¹⁸ “*The position taken in Chomsky (1955/1975) can be supported, I believe, by arguments parallel to those given in Marantz (1984) for the claim that it is VP that assigns a θ -role to the matrix subject, and not simply V.*”¹⁹ Larson (1988, 340).

O tratamento sintático da semântica lexical por meio de uma cadeia de núcleos, posto à luz por Larson, é uma das ideias mais fecundas de toda a teoria gramatical e é um dos fundamentos da sintaxe atual. Esta é a grande contribuição de Larson.

Já a derivação da estrutura de duplo-objeto a partir da de duplo-complemento, por meio de uma transformação, veio a apresentar problemas fundamentais. Análises muito bem fundamentadas, como veremos alguns parágrafos a frente, parecem indicar diferenças semânticas entre as duas estruturas, inviabilizando a derivação de uma a partir da outra.

A solução de Larson à estrutura de duplo-complemento mostrou-se de fundamental importância para a sintaxe, mas pode-se considerar que a estrutura de duplo-objeto permaneceu intocada.

8.7.4 A proposta de Hale e Keyser

Como vimos anteriormente na seção 4.2.3, p. 28, Hale e Keyser propuseram um programa de pesquisa em que a interpretação temática dos argumentos resulta das relações estruturais internas ao constituinte em que eles figuram: “*Argument structure is determined by properties of lexical items, in particular, by the syntactic configurations in which they must appear.*”, Hale & Keyser (2002, p. 1).²⁰ Segundo tal proposta, a teoria temática torna-se redundante como módulo gramatical independente. Para a implementação técnica do seu projeto, H&K ampliaram a proposta sintática de Larson, aplicando-a à projeção argumental de itens lexicais em geral, não apenas à de verbos.

Nesta seção abordaremos certos aspectos da teoria de H&K que dependem de soluções

¹⁸A diferença quanto ao ano de publicação da tese de Marantz (1981 para mim e 1984 para Larson) se deve a que minha citação referencia o texto originalmente publicado.

¹⁹“A posição tomada em Chomsky (1955/1975) pode ser sustentada, creio, por argumentos paralelos aos dados em Marantz (1984) em favor da alegação de que é o VP que atribui um papel- θ ao sujeito da sentença principal, e não simplesmente V.”

²⁰“A estrutura argumental é determinada por propriedades dos itens lexicais, em particular, pelas configurações sintáticas em que eles devem figurar”.

sintáticas ao estilo da concha larsoniana e também analisaremos as representações por eles propostas para as estruturas de duplo-objeto e de duplo-complemento.

Hale e Keyser afastam-se do lexicalismo chomskiano ao propor que os princípios que regem a formação lexical são os mesmos que dirigem a derivação sentencial. Para H&K a estrutura interna de um item lexical é formada sintaticamente: “*The verb cough ... consists of two elements: a root and a verbal nucleus*”, Hale & Keyser (2002, p. 2).²¹ O verbo, assim, é a configuração sintática, composta de uma raiz e de um núcleo. Entretanto H&K distinguem a sintaxe da derivação lexical da sintaxe da derivação sentencial, contrariamente à proposta da Morfologia Distribuída,

Segundo H&K, a fonologia de um constituinte é formada pela operação de *conflation*: “*Conflation consists in the process of copying the p-signature of the complement into the p-signature of the head, where the latter is ‘defective’.*”²² Hale & Keyser (2002, p. 63). Nesta definição, *assinatura-p* significa “assinatura fonológica”, e ela é *defectiva* se é vazia ou se é um afixo, necessitando completar-se com a assinatura-p do complemento. Portanto, *conflation* é um fenômeno fonológico. Deve-se ter em conta que *conflation* é uma operação entre núcleos, uma vez que a “assinatura-p do complemento” é, de fato, a *assinatura-p do núcleo do complemento*. Ou seja, *conflation* é também uma operação que ocorre sintaticamente, provocada por uma cadeia de núcleos (em sentido larsoniano). Além disso, *conflation* é uma operação de cópia, e não de movimento de material fonológico.

Portanto, o programa de pesquisa de Hale e Keyser vai muito além do modelo sintático de Larson.²³ A contribuição fundamental dos dois autores à teoria da linguagem, como entendo, consiste em demonstrar a viabilidade de: (a) determinação dos papéis temáticos a partir de configurações estruturais de constituintes; (b) formação sintática da estrutura interna de constituintes lexicais; (c) formação da fonologia de constituintes por meio de cadeias de núcleos.

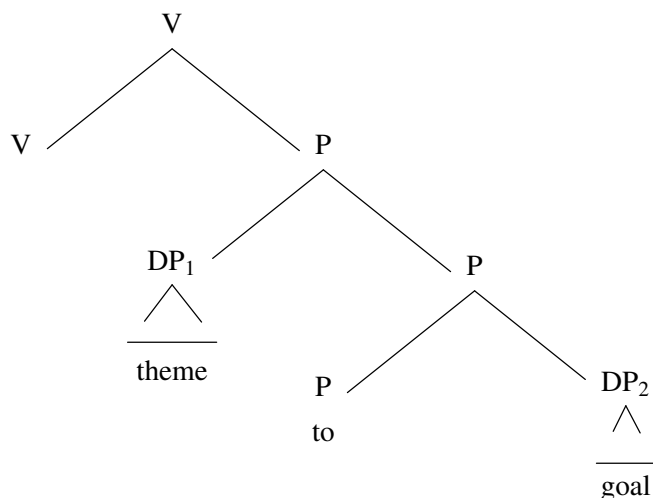
Um caso típico de aplicação do aparato técnico de H&K é a representação da estrutura de duplo-complemento, mostrada no diagrama a seguir (Hale & Keyser 2002, p. 161) .

²¹“O verbo *cough* ... consiste de dois elementos: uma raiz e um núcleo verbal.”

²²“*Conflation* consiste do processo de copiar a assinatura-p do complemento na assinatura-p do núcleo, em que a última [assinatura-p] é ‘defectiva’”

²³Embora o modelo sintático de Larson esteja na base das realizações de H&K. Tem-se a impressão de que Larson não levou o seu modelo às consequências naturais, que seria a formação lexical por meios sintáticos, como o fizeram Hale e Keyser.

(8.25)



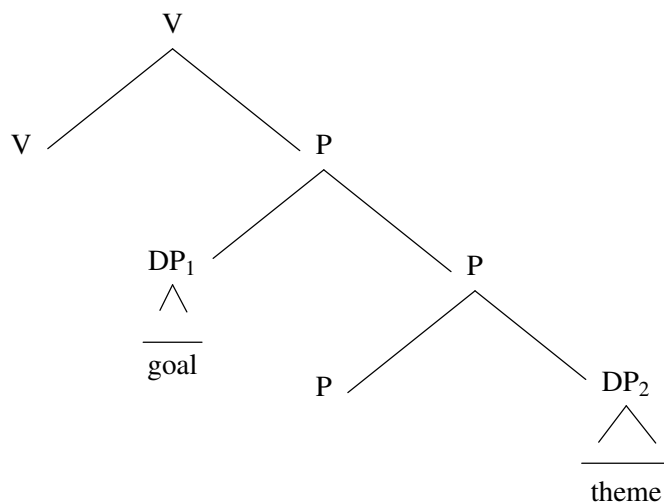
A solução de H&K é isomorfa à de Larson, com uma diferença: o núcleo mais baixo é uma preposição, enquanto que para Larson é V. Este não é um mero detalhe, sob o ponto de vista dos dois autores. Para eles, P é um núcleo que necessariamente projeta a posição de especificador,²⁴ enquanto V projeta apenas complemento. Assim, P aparece sempre em colaboração com V quando se trata da geração de uma estrutura biargumental abaixo de V, que é o caso dos verbos ditransitivos. A verbalização de preposições tornou-se um dos mais importantes recursos da análise sintática após H&K.

Para a geração da estrutura de duplo-objeto, porém, a teoria de H&K enfrenta um contratempo: para obterem-se a sequência pronunciada e as relações de c-comando adequadas entre os argumentos internos, a estrutura verbal deve apresentar a forma exibida no próximo diagrama (Hale & Keyser 2002, p. 160).

²⁴Para H&K, adjetivos, Adj, também projetam especificadores, com o apoio de outros núcleos, mas não nos deteremos neste ponto, no momento.

²⁵Para Larson, PP não é um constituinte que requer um especificador; deste modo, a solução de Larson para a estrutura de duplo-complemento é estritamente binária.

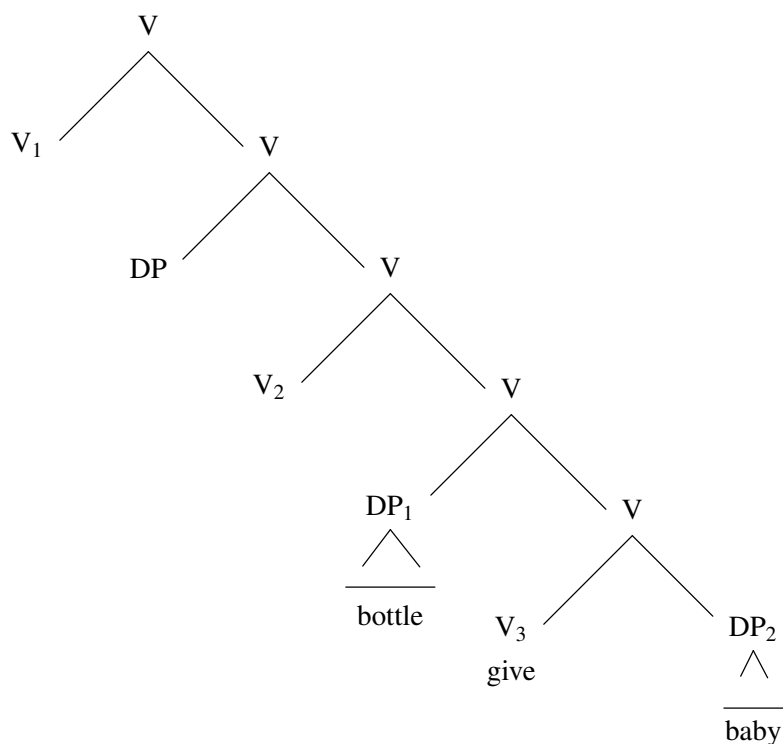
(8.26)



Entretanto, como P é um núcleo fonologicamente nulo, deve necessariamente haver *conflation* entre ele e o núcleo do seu complemento. Este resultado é incorreto para a estrutura de duplo-objeto, em que DP_2 não incorpora.

Tal fato leva H&K a proporem a estrutura do próximo diagrama para a estrutura de duplo-objeto (Hale & Keyser 2002, p. 163).

(8.27)



Nesta solução, V_3 (*give*) estabelece com seus argumentos a mesma relação estrutural existente entre P e seus argumentos, na construção de duplo-complemento. A sequência pronunciada é

obtida após (a) movimento de núcleo, em que *give*, sucessivamente, move-se para V_2 e V_1 , e (b) DP_2 (*baby*) é alçado à posição do especificador de V_2 , “. . . verbo [que] *deve estar presente para permitir* [este] *movimento e lá aparece somente por esta razão*”, p. 164.

A estrutura de duplo-objeto de H&K mostra a grande vitalidade da composição de núcleos proposta por Larson: três núcleos verbais aparecem em sucessão (V_1 , V_2 e V_3). Entretanto, a solução deixa várias questões em aberto: (a) como surgem especificadores para núcleos verbais (que, para H&K, somente projetam complementos)?; (b) de que modo DP_2 (*baby*) pode mover-se para DP? A proposta parece ser extremamente *ad-hoc*.

Deve-se, porém, observar que as soluções de H&K para os dois tipos de verbos ditransitivos (de duplo-objeto e de duplo-complemento) são independentes entre si: não há uma estrutura mais primitiva da qual a outra é derivada por transformação, como propuseram anteriormente Chomsky e Larson. A solução de H&K, assim, é imune à crítica de que as duas estruturas apresentam diferenças semânticas inconciliáveis.

8.7.5 A proposta de Harley

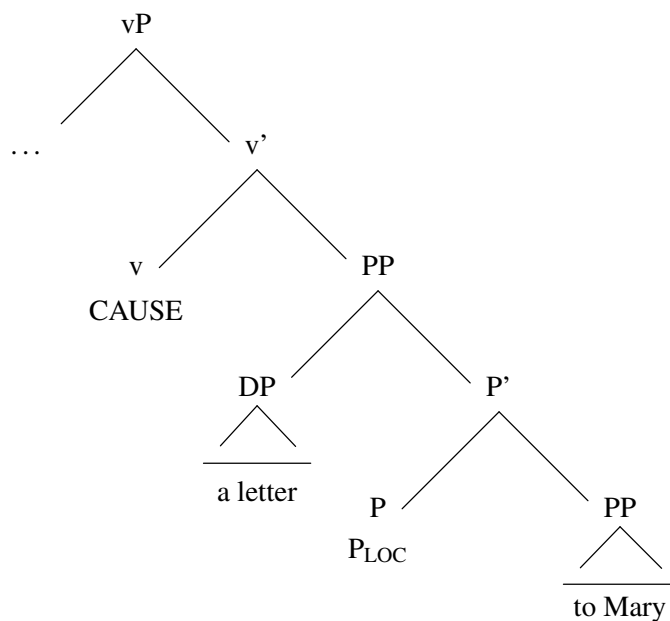
Heidi Harley, em princípio, aceita a representação de Larson para a estrutura de duplo-complemento. Entretanto, analisa as diferenças semânticas existentes entre a estrutura de duplo-objeto e a de duplo-complemento dativa (*to*-dative), mostrando que há uma implicação de posse na primeira, enquanto a segunda apresenta uma interpretação locativa (Harley 2007, 26).

Por exemplo, Harley (2002*b*, 40) observa que a primeira sentença abaixo pode expressar a noção de que Mary está grávida de John, enquanto a segunda implica simplesmente a existência de uma criança que foi fisicamente transferida.

- (9) a. John gave Mary a child
 b. John gave a child to Mary

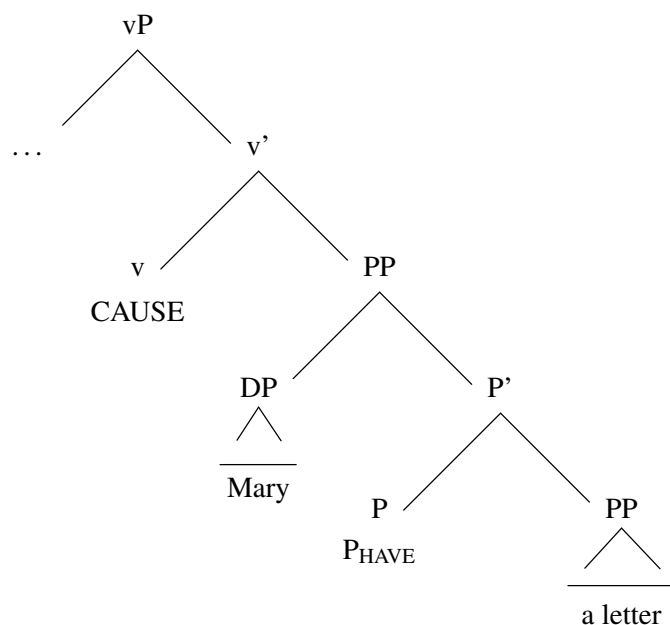
A diferença semântica acarreta não se poder tomar uma das sentenças como básica e a outra como dela derivada por uma transformação. Harley então postula, a partir de uma proposta de Pesetsky, que (a) a estrutura de duplo-objeto é derivada da preposição P_{HAVE} , que dá origem aos verbos *dar* e *ter* e (b) a estrutura de duplo-complemento é derivada da preposição locativa P_{LOC} de acordo com os dois próximos diagramas (ambas as preposições e os verbos delas derivados serão analisados nas seções 8.8.4 e 8.8.3 respectivamente):

(8.28)



Derivação parcial da estrutura de duplo-complemento (Harley 2002*b*, p. 32): '[John] gave a letter to Mary'/'[João] deu um livro a Maria'.

(8.29)



Derivação parcial da estrutura de duplo-objeto (Harley 2002*b*, 32): '[John] gave Mary a letter'/'Mary has a letter'.

Essencialmente, adoto a solução de Harley para as estruturas de duplo-complemento e de duplo-objeto, interpretando-as segundo o paradigma que proponho, como é mostrado nas seções acima citadas.

8.8 Preposições

Passaremos a investigar a estrutura argumental e fonológica dos verbos que derivam de preposições. Estruturalmente, verbos desta classe possuem dois argumentos internos, embora, muitas vezes, tal fato não se mostre visível na sentença pronunciada.

Antes de prosseguir, convém prestar um esclarecimento: o termo *preposição* é usado neste trabalho por tradição, mas verdadeiramente o tema se desenvolve em torno da ideia de *adposição*, um núcleo relacionador de dois *nPs*. Portanto, os identificadores *P* e *p*, que respectivamente denotam um núcleo e o operador projetado por esse núcleo, referenciam a ideia abstrata de *adposição*, embora exibam a evidente conotação de *preposição*. Utilizaremos o mesmo termo, “preposição”, para denotar o núcleo *P* e também uma de suas possíveis manifestações fonológicas.

O núcleo *P* exibe de modo especialmente acentuado uma propriedade comum a todos os núcleos: a fonologia somente se torna visível mediante sua participação em outras estruturas. A manifestação fonológica de *P* é extremamente dependente do contexto sintático em que está inserido. Dependendo do idioma ou da estrutura particular, esse núcleo pode expressar-se por meio de preposições ou posposições, afixos ou ordem de constituintes, ou mesmo de alguma outra forma, todos diferentes possibilidades de um núcleo se manifestar.

De um modo geral, uma estrutura sintática pode apresentar diferentes expressões morfo-sintáticas e, evidentemente, semânticas, dependentes do contexto em que se insere, em função das cadeias de núcleos formadas. O capítulo se desenvolverá sobre este tema.

Basicamente as análises se desenvolverão sobre duas classes de verbos, que constituem duas faces do mesmo fenômeno. Primeiramente abordaremos os verbos gramaticais *dar* e *ter*; em seguida, passaremos aos verbos denominais, como *engavetar*, *entubar* e *atemorizar*. Embora superficialmente os verbos denominais exibam, em português, uma estrutura transitiva “básica”, SVO — ‘O médico entubou o paciente’ —, esses verbos são estruturalmente biargumentais, de acordo a linha teórica da pesquisa.

8.8.1 A estrutura verbal biargumental

Uma *estrutura verbal biargumental* é aquela em que há dois argumentos nominais introduzidos na derivação abaixo do núcleo verbalizador; ou seja, em que o núcleo verbalizador possui dois argumentos nominais “internos”. Estruturas deste tipo incluem as de duplo-objeto, como em ‘John gave Mary a book’, e as de duplo-complemento, como em ‘John gave a book to Mary’/‘João deu um livro a Maria’, estruturas conjuntamente denominadas “ditransitivas”. Mas também incluem os *verbos denominais*, em que um dos argumentos se apresenta de modo menos óbvio, como em ‘Maria engavetou o livro’, em que o argumento *gaveta* aparece em forma verbal. No início do capítulo apresentamos as várias abordagens que contribuíram para o tratamento da estrutura verbal biargumental que passaremos a desenvolver.

8.8.2 Verbos gramaticais biargumentais

Analisaremos agora os verbos gramaticais *dar* (ditransitivo) e *ter* (transitivo), como exemplos de verbos biargumentais.

Os verbos ditransitivos, do mesmo modo que os denominais, resultam da verbalização de estruturas preposicionadas. Mas, contrariamente aos denominais, nos ditransitivos os dois argumentos internos à verbalização são visíveis nas sentenças pronunciadas.²⁶ Nos limites da pesquisa, não é possível fazer uma análise dos verbos ditransitivos em geral; serão, assim, abordados apenas os dois verbos gramaticais. Presumivelmente os métodos de análise a eles aplicados têm validade para os demais ditransitivos.

Historicamente, a análise dos verbos ditransitivos tem como um dos principais eixos as semelhanças e diferenças entre as estruturas de duplo-objeto e de duplo-complemento, respectivamente exemplificadas nas sentenças a seguir:

- (10) a. John gave Mary a book
 b. John gave a book to Mary

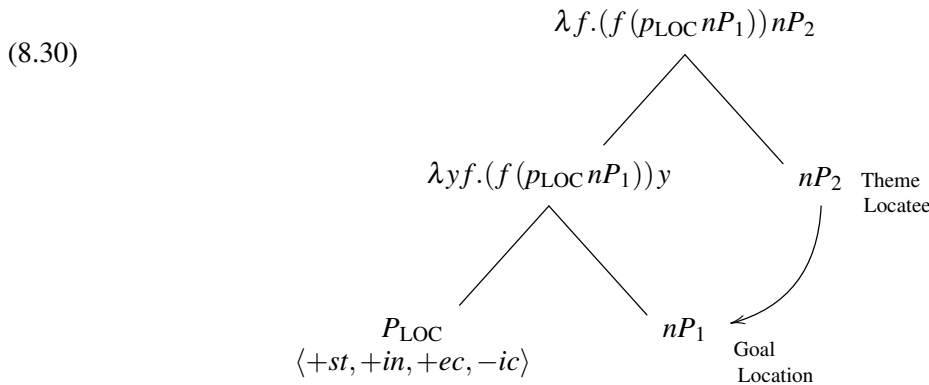
As estruturas ditransitivas foram extensivamente analisadas por Heidi Harley em uma série de artigos (Harley (1998); Harley (2002b); Harley (2007)). Harley representou os verbos *give* (com duplo-objeto) e *have* por meio da preposição P_{HAVE} e o verbo *give* (com duplo-complemento) por meio de P_{LOC} . Baseada em profundas e amplas análises semânticas, a autora mostrou que a estrutura de duplo-objeto implica a ideia de instauração de uma relação de posse, real ou pretendida, e que a estrutura de duplo-complemento significa, essencialmente, uma transferência de lugar. Harley estabeleceu, assim, a semântica fundamental das preposições P_{HAVE} e P_{LOC} .

Em linhas gerais, as conclusões de Harley serão adotadas neste trabalho, ainda que as soluções estruturais que proponho difiram das da autora.

8.8.3 Estrutura de duplo-complemento

Primeiramente, abordaremos a estrutura de duplo-complemento, presente em ‘João deu um livro a Maria’/‘John gave a book to Mary’. Segundo os pressupostos da pesquisa, as estruturas verbais biargumentais são modeladas pela verbalização de estruturas preposicionadas. Proponho, assim, a estrutura representada no diagrama abaixo, como base de geração do verbo *dar/give* da sentença de duplo-complemento.

²⁶O termo “visíveis” está sendo tomado em sentido metafórico.



Há vários pontos a serem justificados nesta estrutura:²⁷

(a) *Análise dos papéis temáticos:*²⁸ A preposição P_{LOC} situa-se na base de geração de 'João deu um livro a Maria'. Como objeto direto, o constituinte *um livro* deve figurar como argumento externo da preposição, e *Maria* como argumento interno. Em uma análise tipicamente gerativista, *um livro* deve c-comandar *Maria*. Assim, tem-se, segundo o diagrama, nP_2 (*um livro*) e nP_1 (*Maria*). Considerada a semântica da preposição, de transferência de local, nP_2 desempenha o papel de Theme/Locatee e nP_1 o de Goal/Location.

(b) *Análise dos traços semânticos:* Pelo fato de tratar-se de uma estrutura relacional, a preposição possui o traço $+in$. A existência de um argumento externo (sujeito) em 'João deu um livro a Maria' torna P_{LOC} compatível com a causalidade externa (ec ou $+ec$). Entretanto, com a preposição em análise, não há como formar uma sentença verbal ativa em que nP_2 (*um livro*) seja o sujeito (interno). Ou seja, não há a possibilidade da causalidade interna, o que proporciona o traço $-ic$. Deste modo, para formação da sentença transitiva, a causalidade externa é obrigatória, fato que implica a seleção do traço $+ec$. A única possibilidade que permanece em aberto é a de se formar uma sentença com sujeito interno por meio da voz média; mas uma análise sumária mostra que uma tal possibilidade também não se realiza. Isto significa que P_{LOC} não pode possuir a conjunção de traços $\langle +ac, +in, +ec \rangle$ (traços da preposição locativa ativa P_{ALOC} , como veremos na seção 8.8.7, p. 144), uma vez que, com esses traços, a formação da voz média seria inevitável. Exclui-se assim a possibilidade do traço $+ac$, pois este traço provocaria a conjunção proibida de traços, considerando-se que os traços $+in$ e $+ec$ são obrigatórios. Tem-se então uma preposição estativa, dotada do traço $+st$, como seria de se esperar de uma preposição locativa. Configuram-se, assim, os traços $\langle +st, +in, +ec, -ic \rangle$ da preposição.

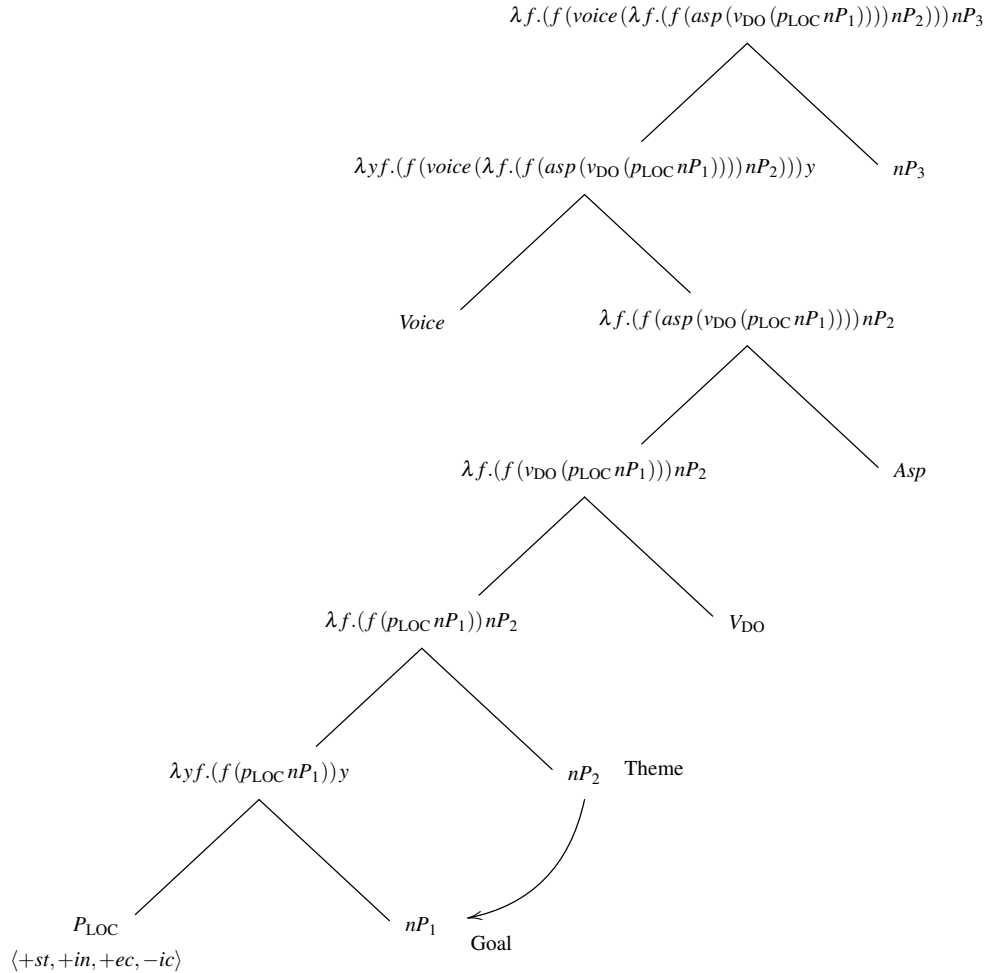
A derivação de 'João deu um livro a Maria'/'John gave a book to Mary' é imediata: o traço $+ec$ de

²⁷No diagrama anterior, e na maioria dos próximos, a seta está direcionada para o argumento Goal (ou análogos), sem maiores motivações semânticas.

²⁸A análise dos papéis temáticos é importante, uma vez que é necessário assegurar a correta interpretação temática dos argumentos em bases puramente estruturais, sem o recurso a grades temáticas ou à codificação semântica de eventos.

P_{LOC} torna a estrutura do diagrama anterior compatível com os núcleos V_{DO} , Asp e $Voice$, este último, introdutor do argumento externo nP_3 , como está representado no próximo diagrama.

(8.31)



Derivação de 'João deu um livro a Maria'/'John gave a book to Mary'.

8.8.4 Estrutura de duplo-objeto

Chegamos, agora à estrutura de duplo-objeto 'John gave Mary a book', que tanta análise motivou no seio do gerativismo chomskiano. Um dos pontos fundamentais, desde a proposta pioneira de Chomsky (1955/1975), é estabelecer se a estrutura de duplo-objeto pode ser derivada, por transformação, da estrutura de duplo complemento de 'John gave a book to Mary'.

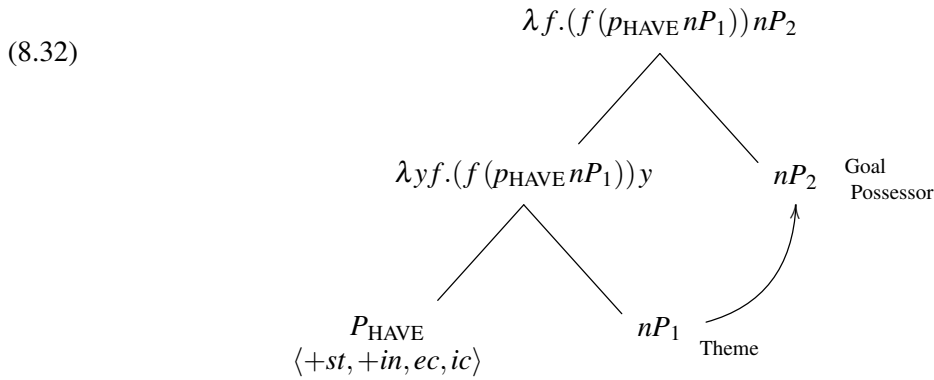
As análises semânticas de Harley mostraram que há diferenças de significado entre as duas classes de sentenças, fato que contradiz a análise derivacional de Larson. Como já comentado anteriormente, Harley propôs que a estrutura de duplo-objeto possui uma semântica de relação de posse, enquanto que a estrutura de duplo-complemento significa uma transferência de lugar. Harley também mostrou a conexão intrínseca entre o verbo *have/ter* e o *give* dotado de duplo objeto: ambos são derivados da preposição P_{HAVE} , a preposição que codifica a relação

de posse.

A geração de *give* e *have* a partir da mesma estrutura soluciona um problema que perpassou um longo período da história do gerativismo chomskiano, o das relações de c-comando entre os dois argumentos do *give* dotado de duplo-objeto.

Adoto as conclusões de Harley, embora a solução que proponho apresente uma implementação distinta.

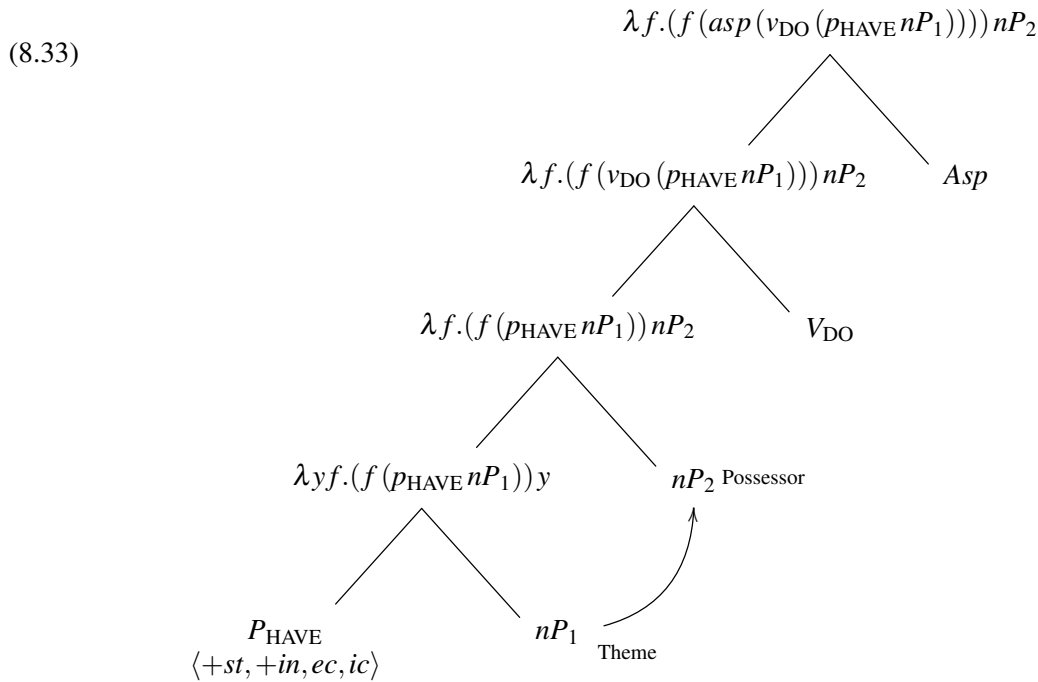
Estrutura proposta para P_{HAVE} .



Justificativa da estrutura de P_{HAVE} :

- (a) *Posicionamento dos argumentos*: A questão é decidir, em 'John gave Mary a book', a relação de c-comando entre os argumentos. A chave da solução encontra-se em 'Mary has a book', derivada da mesma preposição: Mary é um sujeito (interno), e book um objeto; portanto, Mary deve c-comandar book. Considerando-se o diagrama acima, tem-se então: nP_2 (Mary) e nP_1 (book). Mary é o argumento externo da preposição e book, o argumento interno.
- (b) *Papeis temáticos*: a semântica da preposição, proporciona, de modo imediato: nP_2 , Goal/Possessor; nP_1 , Theme.
- (c) *Traços semânticos*: A existência de um sujeito externo em 'John gave Mary a book' e de um sujeito interno em 'Mary has a book' mostra que a preposição é dotada dos traços *ec* e *ic*. Por se tratar de um núcleo relacionador, a preposição possui o traço *+in*. O traço estativo, *+st*, decorre da inexistência de uma estrutura verbal média derivada da preposição. Resultam assim os traços semânticos $\langle +st, +in, ec, ic \rangle$.

Qualquer dos traços semânticos *ec* ou *ic* possibilita a geração da estrutura a seguir.

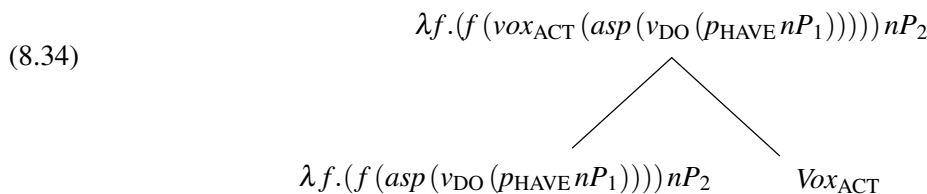


Embora esta derivação possa ser produzida a partir de diferentes traços semânticos, *ec* ou *ic*, cada um dos traços licencia a aplicação de diferentes núcleos ao constituinte situado no topo do diagrama.

O traço semântico *ic*, de causalidade interna, torna o constituinte

$$\lambda f.(f(asp(v_{DO}(p_{HAVE} nP_1)))) nP_2,$$

produzido pela derivação, diretamente compatível com a formação da voz ativa: nP_2 habilita-se como sujeito sentencial interno. Gera-se, assim, a estrutura de 'Mary has a book'.



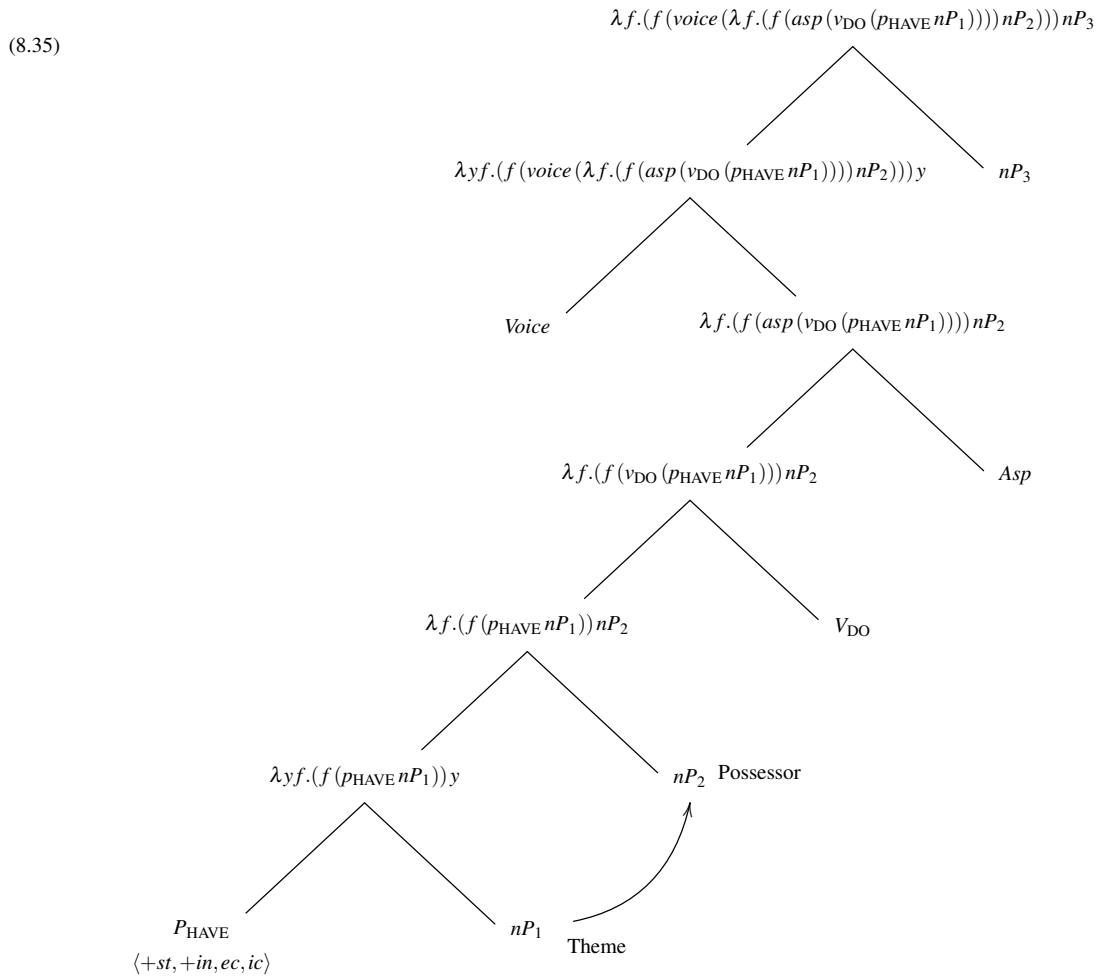
Derivação de 'Mary has a book'.

É importante observar que nP_2 recebe caso como sujeito mesmo tendo sido gerado internamente, abaixo do processo de verbalização, que se inicia com o núcleo V_{DO} . O argumento recebe caso sem que tenha ocorrido qualquer operação de movimento.

Um fato fundamental, sendo válida a estrutura proposta, é o verbo *ter* mostrar-se como um verbo estativo, graças ao traço *+st* de P_{HAVE} , mesmo apresentando uma realização morfo-

fonológica transitiva, característica dos verbos de atividade, fenômeno resultante da verbalização por V_{DO} .

A geração de 'John gave Mary a book' é derivada da aplicação dos núcleos V_{DO} , Asp e $Voice$, licenciados pelo traço *ec* de P_{HAVE} . A derivação é mostrada no próximo diagrama, tendo sido omitidos os núcleos situados acima de $Voice$, inclusive o núcleo de aplicação da voz ativa, situado imediatamente após $Voice$.



Derivação de 'John gave Mary a book'.

8.8.5 Comparação com a solução de Harley

Provavelmente, Harley (2002b) desenvolve a representação conceitualmente mais adequada dos verbos gramaticais *give* e *have*, como também das estruturas de duplo-objeto e de duplo-complemento, segundo a abordagem sintática inaugurada por Hale e Keyser e adotada em nossa pesquisa.

Considero que a solução de Harley (2002b, 32) é preferível à de Harley (2007, 28). A solução deste último texto atribui a *give* um *status* de verbo gramatical na sentença de duplo-

objeto ('Mary gave the car a new engine') e aplica uma interpretação lexical ao mesmo verbo na sentença de duplo-complemento ('Mary gave a new engine to the car'). Por princípio, os verbos biargumentais, no meu modelo computacional, são solucionados por meio de preposições. Deste modo, deixarei à parte as análises mais recentes de Harley (2007) e farei uma comparação entre as propostas do modelo computacional da presente pesquisa e as soluções de Harley (2002b) para *give*, *have* e *get*.

Fundamentalmente, a principal diferença entre as soluções de Harley e as que proponho residem no tratamento aplicado ao operador verbalizador *v*. Para Harley (2002b, pp. 34–61), *v* possui vários “*flavors*”. Os vários modo de atuação de *v*, dependendo da preposição selecionada por este operador, dão origem aos verbos gramaticais por ela analisados: v_{BE} (*have*); v_{CAUSE} (*give*); v_{BECOME} (*get*).

O tratamento de Harley para *v* “sobrecarrega” este operador com funções de natureza bastante diversas: (a) v_{BE} dota *v* da função de um operador categorizador, formador de verbos estativos; (b) v_{CAUSE} atribui a *v* a responsabilidade de um predicado selecionador de argumento externo (“*an external argument-selecting CAUSE predicate*”, p. 63); (c) v_{BECOME} possui a atuação de um núcleo incoativo.

Harley então conclui, com base no polimorfismo do operador *v*, que os múltiplos verbos selecionados (*give*, *have* e *get*) são o resultado de uma visão não lexicalista da sintaxe, possibilitada pela realização tardia do material fonológico (p. 63).

Na abordagem que proponho, o operador *v*, projetado pelo núcleo *V*, apenas atua como operador verbalizador e sob condições bastante restritivas: (a) *v* atua sobre os traços semânticos $\langle +st, +in, -ec, -ic \rangle$, formando a verbalização estativa, situação em que o verbalizador é denotado por v_{BE} ; ou (b) *v* atua sobre os traços *ec*/*+ec* ou *ic*/*–ic*, quando então o verbalizador é denotado por v_{DO} , de verbalização ativa. As diferenças entre os verbos gramaticalmente selecionados resulta das cadeias de operadores formadas, atuantes sobre preposições determinadas, como é mostrado nas derivações deste trabalho. O processamento tardio da fonologia é assegurado pela estrutura geral do núcleo.

Caso as considerações acima estejam corretas, a estrutura de núcleos da presente proposta opera sob condições mais restritas do que as propostas por Harley; cada operador desempenha uma única função e as derivações resultantes obedecem a leis de formação mais bem definidas.

8.8.6 Os verbos denominais

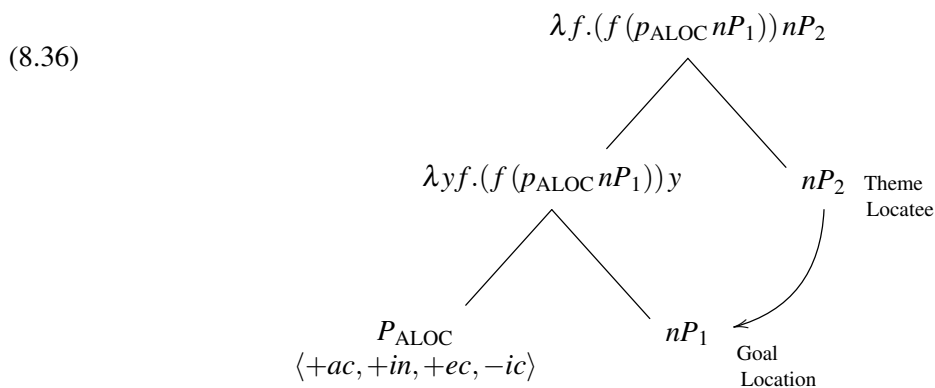
Os verbos denominais manifestam a expressão superficial de um fenômeno gramatical mais profundo, a verbalização de estruturas preposicionadas. Iniciaremos o estudo do fenômeno pela análise de duas classes desses verbos:

- (11) a. Maria engavetou o livro
 b. O político acobertou o escândalo

Embora apareçam sob a forma de verbos transitivos dotados de estrutura argumental canônica (um argumento interno e outro externo) ambos os verbos, segundo o modelo que desenvolvemos, possuem dois argumentos internos, introduzidos por preposições. Esta é a solução proposta em Hale & Keyser (2002) e que se torna inevitável na presente pesquisa, uma vez que não existem verbos derivados de “raízes nominais” ou “verbais”; as raízes são acategoriais quanto à classificação em termos de verbos e nomes.²⁹

8.8.7 A preposição locativa ativa

As preposições estabelecem relações entre dois constituinte nominais. No caso de ‘Maria engavetou o livro’, independentemente do caráter verbal da frase, existe uma *relação de local* entre uma *gaveta* e um *livro*. Postularemos, assim, a existência de uma preposição locativa ativa, P_{ALOC} , que associa os dois nPs , de acordo com a estrutura representada no diagrama abaixo.



No diagrama, *o livro* ocupa a posição do constituinte nP_2 e *gaveta* a de nP_1 . A localização de *gaveta* como argumento interno de P_{ALOC} resulta da morfologia do verbo, *engavetar*, em que nP_1 encontra-se em relação direta com o núcleo. Considerando-se a estrutura da preposição, resta a *o livro* a posição de argumento externo do núcleo. Há outras razões a serem consideradas para a localização dos argumentos, às quais voltaremos posteriormente. Os papéis de Theme/Locatee e Goal/Location, atribuídos respectivamente a nP_2 e nP_1 decorrem, de modo, natural, da semântica da preposição.

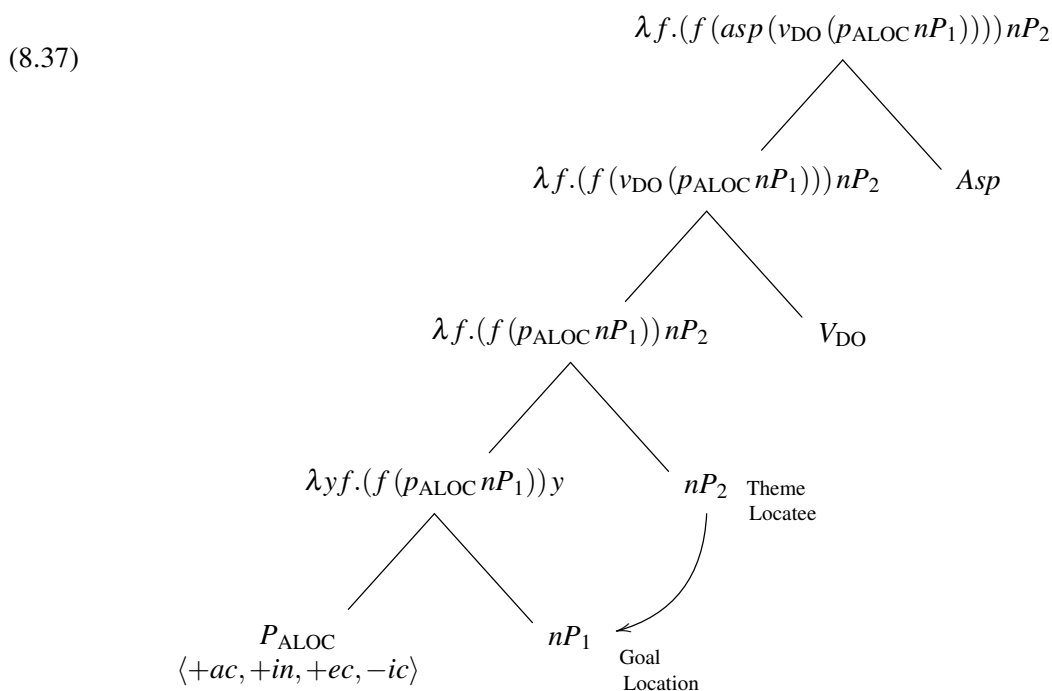
Os traços semânticos $\langle +ac, +in, +ec, -ic \rangle$ indicam os tipos de causalidade compatíveis com o núcleo. No caso, o traço $-ic$ sinaliza que não se pode gerar uma sentença como ‘O livro engavetou’, em que o sujeito tem origem interna; isto é, o argumento nP_2 não pode ser sujeito de uma sentença ativa. Já o traço $+ec$ revela a compatibilidade com um sujeito externo, possibilitando a derivação de sentenças como ‘Maria engavetou o livro’. Há ainda a possibilidade de sentenças como ‘O macaquinho engavetou-se’, em que o sujeito tem origem interna, e que será

²⁹ As raízes são lexicalmente categorizadas como *raízes*.

analisada brevemente. Se existem restrições quanto ao tipo de sujeito sentencialmente aceitável ('A caneta engavetou o livro', 'O livro engavetou-se'), devem-se a propriedades semânticas não contempladas nos traços selecionados para o núcleo e não a características estruturais dos ambientes verbais.

A compatibilidade com a formação da voz média dota a preposição do traço de atividade $+ac$. O traço de incidência, $+in$, é obrigatório para os núcleos relacionadores, indicando que a *composição* do núcleo com o seu argumento interno (no caso, $P_{ALOC} nP_1$) requer um argumento, nP_2 .³⁰ Existe uma outra possibilidade, em que o traço de incidência é selecionado com o valor *in*, de não-necessidade. Neste caso, a preposição seleciona necessariamente o argumento interno, e, opcionalmente, o externo. É possível, então, gerar expressões como *em casa*, sem o argumento externo. Até o momento, nas derivações que investiguei, uma tal possibilidade não ocorreu. De todo modo, é uma questão empírica determinar o valor do traço, segundo o modelo adotado; o valor $+in$ foi selecionado por ser mais restritivo.

O traço de causalidade *+ec* torna a estrutura gerada por P_{ALOC} , exibida no diagrama anterior, compatível com a verbalização ativa, por V_{DO} , produzindo a derivação do diagrama abaixo.



Esta estrutura dá origem à sentença transitiva 'Maria engavetou o livro' e à sentença média 'O macaquinho engavetou-se', como veremos a seguir.

³⁰Observar que nP_2 é argumento *interno* de $(P_{\text{ALOC}} nP_1)$ e *externo* de P_{ALOC} .

8.8.8 Derivação da sentença média

Para a geração da sentença média, é necessário ter em conta que o traço $-ic$ impede que nP_2 do constituinte resultante da derivação,

$$\lambda f.(f(asp(v_{DO}(p_{ALOC} nP_1)))) nP_2,$$

visível no topo do diagrama, seja interpretado como sujeito. Isto é, a variável f não pode receber, por exemplo, operadores responsáveis por estabelecer concordância, atribuir caso e *tense*, todos formadores sentenciais. Fica, deste modo, impedida a geração de sentenças como ‘O livro engavetou’, em que o sujeito tem origem interna, abaixo do núcleo verbalizador.

Entretanto, a combinação de traços $\langle +ac, +in, +ec \rangle$ faz com que o constituinte resultante da derivação seja compatível com a voz média, na qual um constituinte nominal internamente gerado, nP_2 , pode ser interpretado como sujeito. A voz média possibilita a geração da sentença ‘O macaquinho engavetou-se’, como é mostrado no próximo diagrama.

$$(8.38) \quad \begin{array}{c} \lambda f.(f(vox_{MID}(asp(v_{DO}(p_{ALOC} nP_1)))) nP_2 \\ \swarrow \quad \searrow \\ \lambda f.(f(asp(v_{DO}(p_{ALOC} nP_1)))) nP_2 \quad \quad \quad vox_{MID} \end{array}$$

A estrutura da sentença média é dada pelo constituinte assim derivado,

$$\lambda f.(f(vox_{MID}(asp(v_{DO}(p_{ALOC} nP_1)))) nP_2,$$

em que:

- (a) nP_2 (O macaquinho), introduzido na derivação abaixo do núcleo verbalizador, ocupa a posição estrutural de sujeito, como argumento externo da cadeia de operadores

$$vox_{MID} \circ asp \circ v_{DO} \circ p_{ALOC}.$$

Pode-se caracterizar informalmente a posição do sujeito como “a do constituinte (nominal) que ocupa a posição extrema-direita de uma derivação verbal”.

- (b) O verbo *engavetou* é formado pela operação de *conflation*, resultante da aplicação do operador p_{ALOC} ao constituinte nP_1 , em que a fonologia *em* do operador se concatena com a fonologia de nP_1 , *gavet*,³¹ produzindo a sequência fonológica *engavet*. A esta sequência,

³¹ A fonologia de nP_1 , neste estágio da derivação, encontra-se isenta de qualquer tratamento fonológico que não seja o resultante da nominalização da raiz (isto é, encontra-se isenta de marcas de concordância, caso, etc.). No exemplo, a fonologia de nP_1 reduz-se a *gavet*.

o verbalizador v_{DO} sufixa a vogal temática, *a*, gerando o resultado fonológico *engaveta*. A fonologia final do verbo pronunciado é formada por núcleos não mostrados na derivação, entre os quais *Tense*.

(c) O clítico *se*, de *engavetou-se*, é a expressão fonológica do operador vox_{MID} (ou seja, é a expressão fonológica da voz média). No caso de um verbo de atividade com sujeito animado, a expressão da voz média, em português, tem interpretação reflexiva. A voz média licencia um argumento interno ao processo de verbalização como sujeito estrutural, sem que para o licenciamento tenha ocorrido movimento do argumento.

8.8.9 O verbo gramatical *pôr*

A análise do verbo *engaveta[r]* proporciona uma visão do processo de formação verbal como expressão fonológica de uma cadeia de operadores aplicada ao seu argumento interno. No caso, é a expressão da cadeia $f \circ vox_{MID} \circ asp \circ v_{DO} \circ p_{ALOC}$ aplicada a nP_1 : a cadeia de operadores contribui com a expressão fonológica gramatical (*en-*) e nP_1 com a expressão fonológica lexical (*gavet-*).

$$(8.39) \quad \begin{array}{c} (f \circ vox_{MID} \circ asp \circ v_{DO} \circ p_{ALOC} nP_1) nP_2 \\ \swarrow \quad \searrow \\ f \circ vox_{MID} \circ asp \circ v_{DO} \circ p_{ALOC} \quad nP_1 \end{array}$$

Há entretanto a situação em que a cadeia de operadores não entra em *conflation* com a fonologia do seu argumento interno nP_1 ; em outros termos, não se apropria da fonologia de nP_1 . Neste caso a cadeia $f \circ vox_{MID} \circ asp \circ v_{DO} \circ p_{ALOC}$ é forçada a manifestar uma expressão fonológica própria, necessariamente gramatical, uma vez que não há operadores lexicais envolvidos. Em português tal fonologia é expressa pelo (paradigma do) verbo *pôr*.

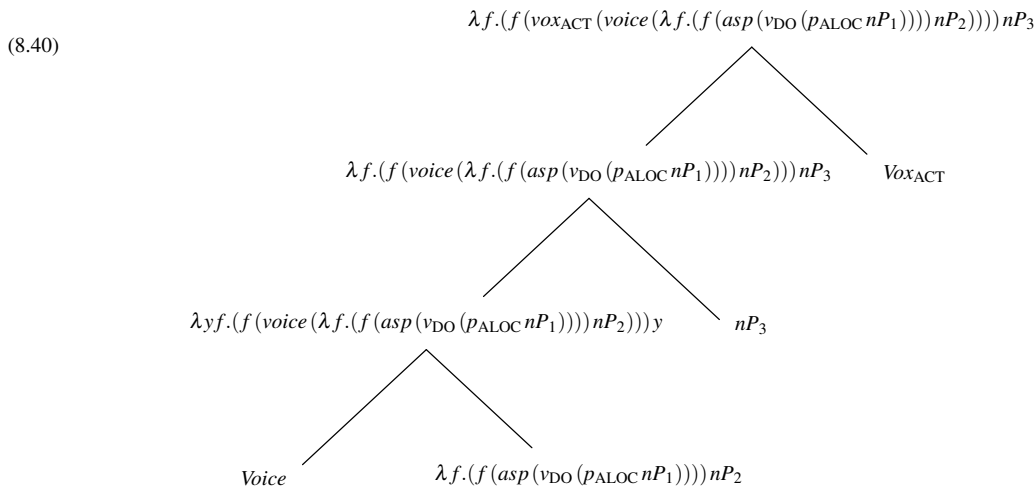
Outro fato porém precisa ser observado: como nP_2 ('livro') ocupa a posição estrutural de objeto da estrutura *gramatical* transitiva ('Maria pôs o livro . . . gaveta'), nP_1 ('gaveta') deve se manifestar *gramaticalmente* como argumento oblíquo ('Maria pôs o livro na gaveta'). A cadeia de operadores contribui então com sua única fonologia específica (a fonologia *default* de p_{ALOC} , *em*) para marcar nP_1 como o argumento complementar ao objeto (marcar como oblíquo). Ou seja, nP_1 sempre se expressará fonologicamente em combinação com a cadeia de operadores, seja por *conflation* (*engavetar*), seja precedido de uma preposição (*em gaveta*).

8.8.10 Derivação da sentença transitiva

A estrutura denominal transitiva ativa é derivada a partir do mesmo constituinte,

$$\lambda f.(f(asp(v_{DO}(p_{ALOC} nP_1)))) nP_2,$$

do qual também se deriva a estrutura média. O traço *+ec*, de P_{ALOC} , torna o constituinte compatível com um argumento externo, nP_3 , o qual é introduzido pelo núcleo *Voice*; a caracterização de nP_3 como sujeito estrutural é feita pelo núcleo Vox_{ACT} , formador da voz ativa, também licenciado por *+ec*.



Tem-se, assim, a derivação da estrutura de ‘Maria engavetou o livro’.

(a) O constituinte nominal nP_3 (*Maria*) ocupa a *posição estrutural de sujeito*, como argumento externo da cadeia de operadores $vox_{ACT} \circ voice$. É importante observar que o sujeito é *estruturalmente* caracterizado por esta cadeia de operadores, independente de qualquer outro operador que venha a ser recebido por meio da variável f . A composição com outros operadores (como atribuidores de *caso* ou *concordância*) produz efeitos fonológicos ou de ordem de constituintes, mas nada acrescentam à caracterização estrutural do sujeito.

(b) nP_2 (*o livro*) ocupa a *posição estrutural de objeto*, como argumento externo da cadeia de operadores $asp \circ v_{DO} \circ p_{ALOC}$, “sob a condição” de esta cadeia também se situar como operador do argumento interno de $vox_{ACT} \circ voice$.³² Aplicam-se a nP_2 as mesmas observações feitas para o sujeito: a caracterização estrutural do objeto independe de operadores externos ao constituinte.

(c) O verbo, *engavetou*, é derivado segundo o processo já analisado para a derivação da sentença em voz média.

³²De acordo com análises anteriores, nP_2 pode também ocupar a posição de sujeito, em um ambiente de voz média.

8.8.11 A preposição condicional

Analisaremos agora as estruturas derivadas de uma preposição mais abstrata do que a anterior. A preposição P_{COND} estabelece uma relação em que nP_2 se encontra sob uma certa condição decorrente de nP_1 .

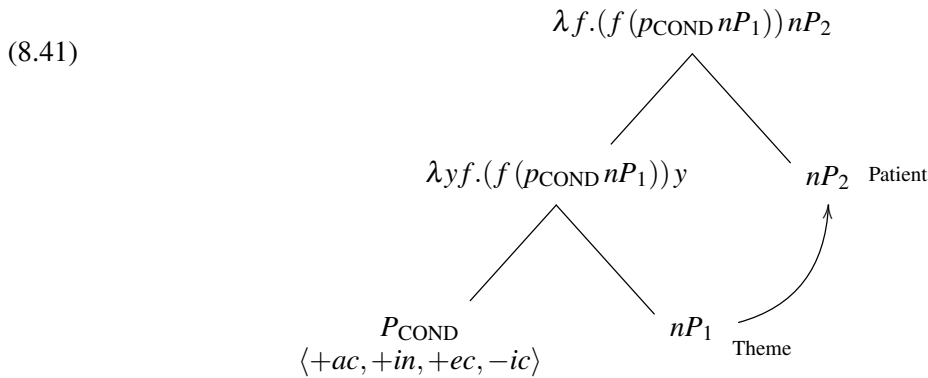
Um exemplo de uso desta preposição encontra-se na sentença

- (12) a. O político acobertou o escândalo

em que claramente o significado está ligado a uma condição aplicada a alguma eventualidade, e não à aplicação de uma coberta a um escândalo, exceto se em sentido figurado. Tem-se, assim, a semântica de algo submetido a uma condição. Tal interpretação é visível em ‘João apimentou a conversa’, em que seguramente não se aplicou, fisicamente, pimenta a uma conversa, e também em ‘João atucanou Maria’, em que é difícil perceber a participação de um tucano no referido fato. As considerações de justificação dos traços semânticos de P_{COND} ,

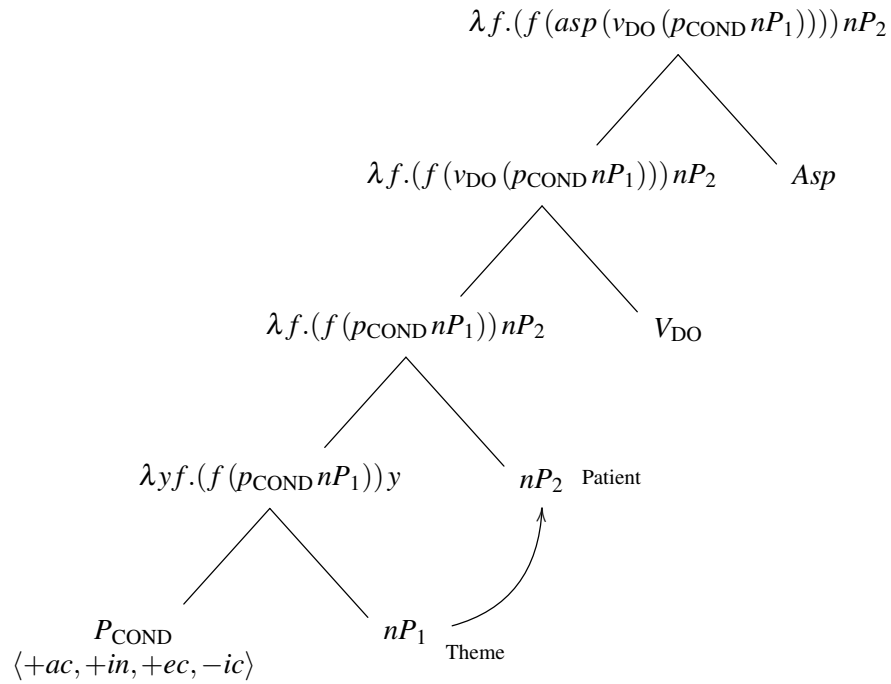
$$\langle +ac, +in, +ec, -ic \rangle$$

são análogas às feitas para P_{LOC} . Do mesmo modo, a derivação das estruturas transitiva (‘João apimentou a conversa’) e média (‘A conversa apimentou-se’) segue exatamente as mesmas linhas das derivações desenvolvidas na seção anterior, para o verbo *engavetar*. A análise dessas derivações não será repetida aqui, apenas os diagramas serão apresentados, para exibir a formação estrutural dos constituintes.



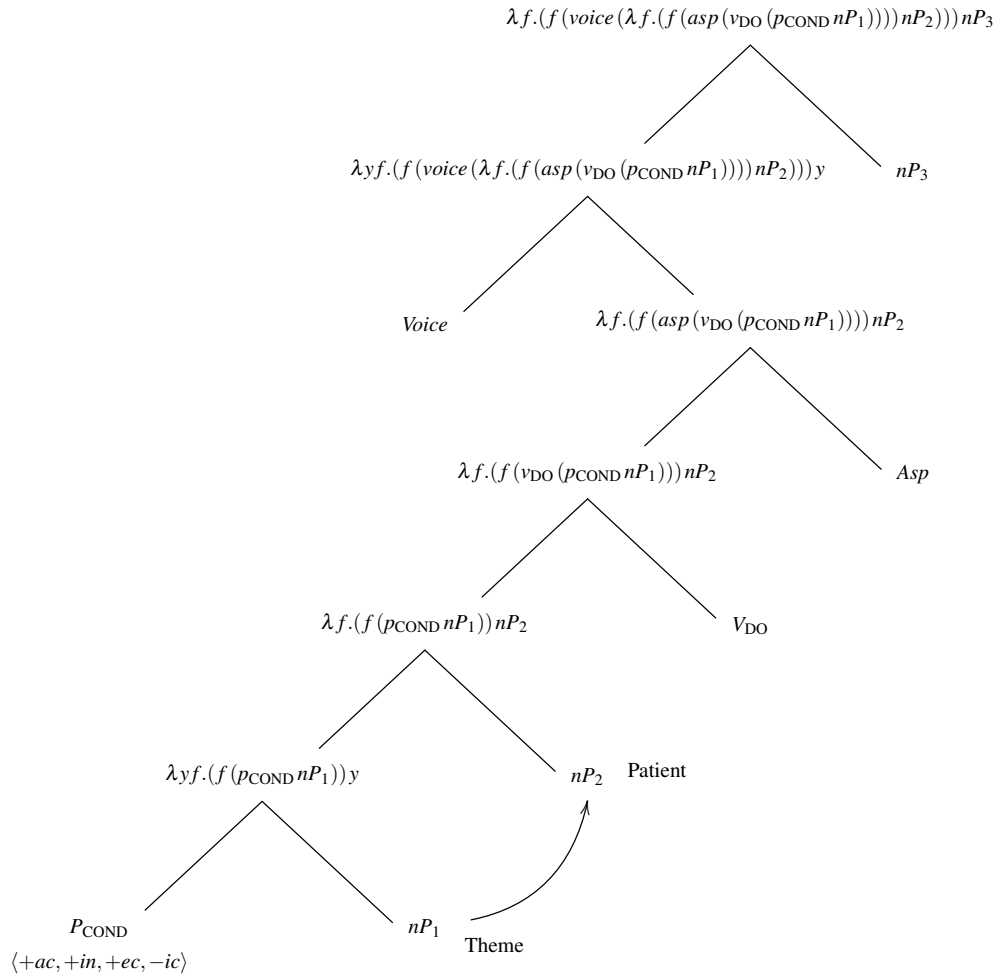
A interpretação temática dos argumentos da preposição condicional precisa ser justificada: (a) a semântica fundamental de P_{COND} não se refere à aplicação física do ente denotado por nP_1 ao ente denotado por nP_2 ; mas (b) a uma condição adquirida pelo último ente, resultante da aplicação de P_{COND} a nP_1 . Deste modo, nP_1 desempenha o papel de Theme da relação e nP_2 o de Patient.

(8.42)



O diagrama acima exhibe a estrutura verbal básica, que dá origem à derivação das sentenças transitivas (como ‘O político acobertou o escândalo’), mostrada no próximo diagrama, e à derivação das sentenças médias, como ‘A conversa apimentou-se’. Como já se observou, os processos de derivação são análogos aos da seção anterior.

(8.43)



8.8.12 Preposição estativa básica

Para finalizar o estudo das preposições, analisaremos a preposição que estabelece a relação genitiva entre dois nPs , denotada por P_{GEN} . O principal método de análise da semântica gramatical de um núcleo é considerar a família de estruturas de que ele pode participar. No caso de P_{GEN} , somente estão disponíveis a nominalização e a verbalização estativa, como nas sentenças a seguir.

(13) Casa de Maria

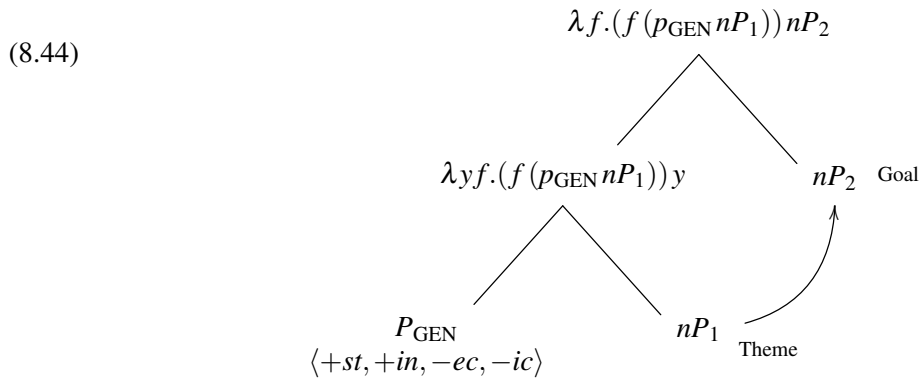
(14) A casa é de Maria

A compatibilidade de um constituinte com V_{BE} , pela minha proposta, resulta dos traços semânticos

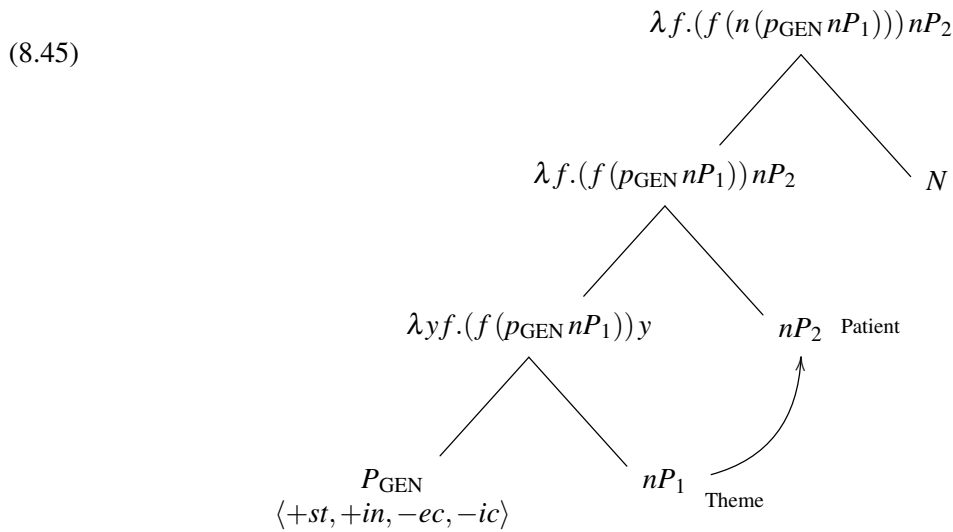
$$\langle +st, +in, -ec, -ic \rangle.$$

Estes então são os traços semânticos de P_{GEN} , estruturalmente representados no diagrama

abaixo.



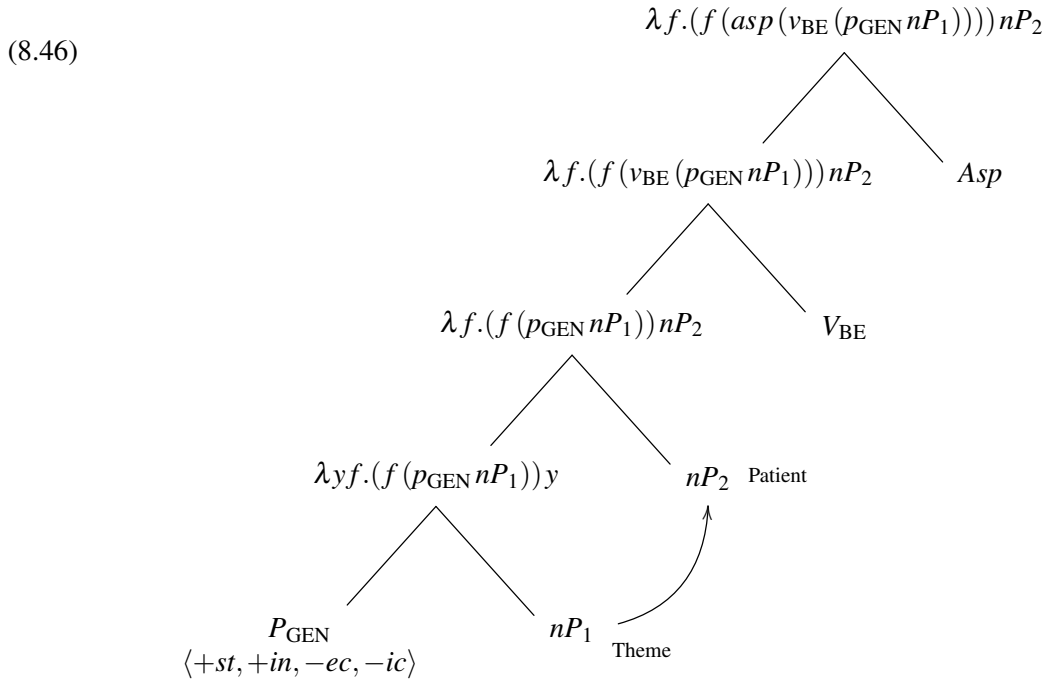
A estrutura pode ser diretamente nominalizada, dando origem ao constituinte ‘Casa de Maria’ cuja derivação está representada a seguir.



Os papéis temáticos atribuídos a nP_1 e nP_2 explicam-se pela analogia existente entre ‘Casa de Maria’ e ‘Casa mariana’: a aplicação da preposição ao seu argumento interno mostra um valor adjetival, valendo, assim, a especificação temática já analisada para a estrutura de estatividade básica.

Na estrutura adjetival (lexical), os traços semânticos de estatividade necessária destituída de causalidade, $\langle +st, +in, -ec, -ic \rangle$, devem-se às propriedades lexicais da raiz; na estrutura genitiva a mesma configuração de traços resulta da aplicação de um núcleo, P_{GEN} .

A sentença ‘A casa é de Maria’ é gerada por verbalização estativa de P_{GEN} , como é mostrado na derivação abaixo.



Nesta estrutura, nP_2 , o argumento externo da cadeia de operadores $asp \circ v_{BE} \circ p_{GEN}$ ocupa a posição estrutural de sujeito. Em português, o operador v_{BE} manifesta-se fonologicamente de modo não defectivo; no caso, v_{BE} se expressa pelo verbo 'é'.³³ O operador p_{GEN} , não se integrando fonologicamente por *conflation* a v_{BE} , é forçado a se manifestar autonomamente, e o faz por meio de uma preposição, 'de'. Deriva-se assim a sentença 'A casa é de Maria'.

8.9 Voz

Abordamos anteriormente as noções de voz ativa e média ao examinar a estrutura inacusativa (seção 8.3) e a derivação dos verbos denominais (seção 8.8.8), como parte da análise de formação do sujeito daquelas estruturas. Uma apresentação mais completa do tema não poderia ser feita então, por depender de elementos das estruturas verbais que ainda viriam a ser desenvolvidos. Investigaremos agora as vozes *ativa*, *média* e *passiva* de modo mais abrangente.

Primeiramente é preciso observar que o núcleo *Vox* categoriza constituintes de estrutura

$$\lambda f.f(\dots v_{DO} \dots)nP,$$

em que $(\dots v_{DO} \dots)$ é o argumento interno e nP o argumento externo da variável ligada f . Dependendo do valor atribuído a f , o nP poderá ocupar a posição estrutural de sujeito ou de objeto, de acordo com os resultados dos capítulos anteriores.

³³Ter em conta que a cadeia $asp \circ v_{BE} \circ p_{GEN}$, na geração da expressão pronunciada, se encontra em composição com outros operadores, como *Tense*, omitidos da derivação.

No caso em que nP é introduzido na derivação abaixo de v_{DO} (é argumento interno de v_{DO}), o constituinte apresenta a estrutura

$$\lambda f.(f(asp(v_{DO}(op\ xP))))\ nP,$$

em que op é um operador ($comp$ ou $prep$) e xP é um constituinte lexical (rP ou nP).

Esta estrutura pode se realizar de diferentes modos, conforme a seguinte exemplificação (nos exemplos o argumento externo é posto entre colchetes, por não pertencer à estrutura): estrutura inacusativa ('Maria caiu'); alternante-ativa ('O rádio quebrou'); alternante-estativa ('O café esfriou'); de duplo-complemento ('[João] deu um livro a Maria'); de duplo-objeto ('Maria tem um livro'); denominal locativa ('O macaquinho engavetou-se'); denominal condicional ('A conversa apimentou-se'); transitiva lexical ('[Maria] lava a criança').

No caso em que nP entra na derivação acima de v_{DO} (é argumento externo de v_{DO}), o constituinte possui a estrutura

$$\lambda f.(f(voice(\lambda f.(f(asp(v_{DO}(op\ xP))))\ nP_1)))\ nP,$$

em que op é um operador (r , $comp$ ou $prep$) e xP é um constituinte lexical (veP , rP ou nP).

Como no caso anterior, há diferentes maneiras de a estrutura se realizar, ilustradas a seguir: estrutura inergativa ('Maria riu'); alternante-ativa ('Maria quebrou o rádio'); alternante-estativa ('Maria esfriou o café'); de duplo-complemento ('João deu um livro a Maria'); de duplo-objeto ('John gave Mary a book'); denominal locativa ('Maria engavetou o livro'); denominal condicional ('João apimentou a conversa'); transitiva lexical ('Maria lava a criança').

Os exemplos mostram que o operador v_{DO} é insuficiente para caracterizar o nP das estruturas acima como sujeito ou como objeto. Segundo as análises efetuadas anteriormente, v_{DO} é mesmo insuficiente para atribuir às estruturas um caráter especificamente verbal: uma estrutura verbal inacusativa, por exemplo, pode ser nominalizada, derivando o nome 'carta caída', como se observa na p. 100.

Tais fatos se devem a que, no constituinte $\lambda f.f(\dots v_{DO} \dots) nP$, a relação definida pela variável f entre nP e $(\dots v_{DO} \dots)$ se encontra indefinida. A cadeia de operadores somente estabelece contato com nP mediante a participação de f e f é um termo não-especificado, assim a relação é não-especificada. Todo o capítulo 5, **Formalização da noção de núcleo**, desenvolve a demonstração deste fato. Portanto, é necessário atribuir a f um operador que interprete como verbal a relação entre seus dois argumentos; este operador é o núcleo Vox .

O núcleo Vox qualifica a relação entre $(\dots v_{DO} \dots)$ e nP como verbal e em consequência caracteriza o nP como sujeito, por sua posição estrutural.

8.9.1 Voz ativa

A semântica fundamental da voz ativa (como vimos na p. 100) caracteriza-se pelo fato de o sujeito ser necessário à iniciação do processo ou estado indicado pelo verbo. Uma vez que os traços de causalidade (*ec* e *ic*) estão ligados à ideia de um ato de iniciação, quaisquer desses traços licencia a aplicação de Vox_{ACT} , ou seja, o núcleo Vox na modalidade ativa.

No caso dos constituintes de estrutura

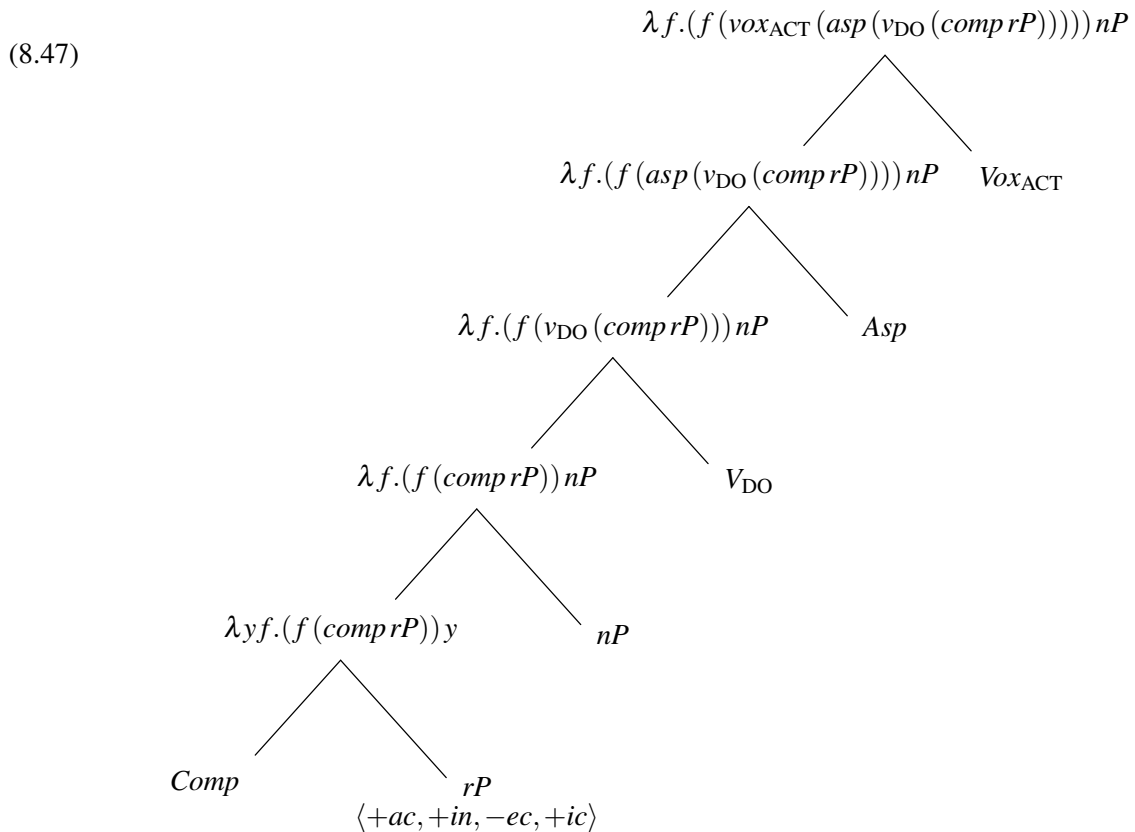
$$\lambda f.(f(asp(v_{DO}(opxP))))nP,$$

devido à origem interna do nP , o licenciamento de Vox_{ACT} somente é possível se a semântica de $(asp\ v_{DO}(opxP))$ contiver os traços de causalidade interna *ic* ou *+ic*: o traço *-ic* impede a caracterização do nP como sujeito e, por consequência, o licenciamento de Vox_{ACT} .

Os traços licenciadores da voz ativa por constituintes dotados de argumentos internos são assim

$$\langle +ac / +st, +in, \dots, ic / +ic \rangle,$$

em que \dots indica que a configuração semântica independe dos traços *ec*. Uma derivação típica em que Vox_{ACT} é licenciada por essa composição de traços é mostrada no diagrama a seguir.



Licenciamento de Vox_{ACT} : derivação verbal inacusativa.

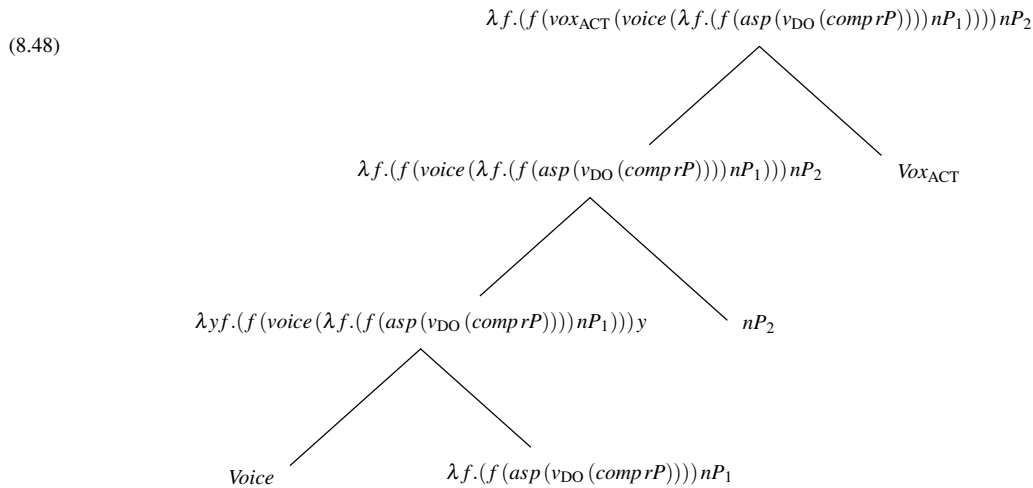
Os constituintes de estrutura

$$\lambda f.(f(\text{voice}(\lambda f.(f(\text{asp}(v_{\text{DO}}(\text{op}xP))))nP_1)))nP,$$

em que nP entra na derivação acima de v_{DO} , são sempre compatíveis com a voz ativa, uma vez que a introdução do núcleo *Voice* pressupõe um dos traços *ec* ou *+ec*. Como o licenciamento de Vox_{ACT} por este tipo de constituintes independe dos traços de incidência e de causalidade interna, os traços de licenciamento são portanto

$$\langle +ac / +st, \dots, ec / +ec, \dots \rangle.$$

O próximo diagrama exibe uma derivação representativa da configuração de traços categorizada por Vox_{ACT} .



Licenciamento de Vox_{ACT} : verbo transitivo.

As condições de licenciamento mostram que a voz ativa não se restringe a verbos transitivos e “de ação”. É perfeitamente possível um argumento interno de estrutura estativa receber o tratamento gramatical de sujeito ativo, como se depreende de ‘O repolho azedou’; também é possível a existência de um verbo transitivo ativo derivado de uma estrutura estativa em que todos os argumentos são internos, como é o caso do verbo *ter*.

8.9.2 Vozes média e passiva

A semântica fundamental da voz média, segundo Benveniste (1950a, 187), é aquela em que “o verbo indica um processo do qual o sujeito é a sede; o sujeito está no interior do processo”. Adotaremos esta semântica nas análises a seguir. Nos termos da presente pesquisa, o processo é indicado por um verbo necessariamente dotado de argumento interno (possuidor do traço de incidência *+in*) na condição de sujeito. A voz média é portanto uma operação (que

denotaremos pelo núcleo Vox_{MID}) de licenciamento de um argumento interno como sujeito. Nada se afirma entretanto a respeito da causalidade do processo, se externa ou interna; apenas se diz que o sujeito está no interior do processo.

Em português, a voz média origina sentenças como 'Maria operou-se' e 'João barbeou-se', ambas compatíveis com a semântica descrita. O núcleo Vox_{MID} deve então impedir a geração de sentenças como 'Maria caiu-se'.³⁴

Investiguemos a composição de traços licenciadores de Vox_{MID} segundo a conceituação acima.

- (a) Até o momento, os únicos traços que se podem tomar como assegurados são $+ac/+st$ e $+in$, sendo necessário analisar então os traços de causalidade.
- (b) Deve-se excluir a possibilidade de geração de uma sentença como 'Maria caiu-se', fato que exclui $-ec$ da configuração possível de traços; a impossibilidade de $-ec$ implica a impossibilidade do traço $+ic$. Ou seja, eliminam-se de imediato os traços característicos da estrutura inacusativa.
- (c) É possível entretanto 'Maria cortou-se', mas é agramatical 'Maria cortou', em que 'Maria' é um sujeito interno. Ou seja, a voz média pode formar 'Maria cortou-se' mesmo em presença do traço $-ic$ das construções transitivas. Tal fato significa que a formação da voz média não pressupõe a necessidade da causalidade interna, que pode não existir.

Temos assim a configuração de traços licenciadores do núcleo Vox_{MID} :

$$\langle +ac/+st, +in, +ec/ec, -ic/ic \rangle.$$

A configuração de traços revela um núcleo que efetivamente *pressupõe* e existência de causalidade externa, sem a necessidade da causalidade interna, mesmo contendo o traço $+in$ de presença obrigatória de argumento interno. As análises do restante do capítulo mostram que a voz média licencia um argumento interno como sujeito, desde que este sujeito receba *uma interpretação de sujeito externo*. Tal afirmação é mais forte do que a simples suposição de existência de causalidade externa para um sujeito internamente gerado.

Passemos agora à investigação dos traços semânticos da voz passiva, que denotaremos por Vox_{PAS} . De certo modo, a análise é mais simples do que a empreendida para a voz média:

- (a) a voz passiva se forma mesmo diante da impossibilidade de causalidade interna ('A bola foi chutada');
- (b) pressupondo porém a existência de um argumento interno; e
- (c) também reque-

³⁴Não se está, com isso, afirmando que a sentença 'Maria caiu-se' não pode ser gerada pelo sistema da língua, apenas que não é gerada pelo sistema de formação argumental. Em uma sentença como 'Maria caiu-se de amor pelo Pedro' (sugestão do Pagani) entram em jogo os argumentos aplicados, presentes também em sentenças como 'Não é que o piá me quebrou a vidraça da velha?'. Os argumentos aplicados não fazem parte propriamente da estrutura argumental lexical (no exemplo anterior fica claro que o enunciador não foi quebrado pelo referido piá), constituindo fenômenos que atuam sobre a estrutura argumental já formada. Durante a nossa pesquisa, dedicamos muitas análises aos argumentos aplicados, sobretudo sob o ponto de vista estrutural. Por limitações de tempo e de foco do trabalho a divulgação dessas análises foram deixadas para ocasião futura.

rendo causalidade externa, $+ec$ ou ec . Tem-se portanto exatamente a mesma configuração de traços que a existente para a voz média:

$$\langle +ac / +st, +in, +ec / ec, -ic / ic \rangle.$$

A configuração de traços acima, licenciadora de Vox_{MID} e Vox_{PAS} , é muito interessante, por incluir a combinação de traços ec e ic que possibilita a alternância de transitividade das estruturas ativas. Temos então que os traços que caracterizam a variante intransitiva das estruturas alternantes (como ‘O repolho azedou’ e ‘Maria tem um livro’)³⁵ constituem um subconjunto dos traços licenciadores das vozes média e passiva. Ou seja, os traços que licenciam Vox_{MID} e Vox_{PAS} também licenciam uma modalidade especial de Vox_{ACT} , a formadora de estruturas verbais ativas alternantes (dotadas de sujeito interno).

Portanto, as vozes média e passiva (e parcialmente a ativa) são licenciadas pela estrutura dotada de argumento interno e por idênticos traços semânticos, como é mostrado no próximo diagrama (em que Op é um dos núcleos $Comp$ ou $Prep$ e xP é rP ou nP).

Tem-se então que as vozes média e passiva, e parcialmente a ativa, se caracterizam pelos traços semânticos

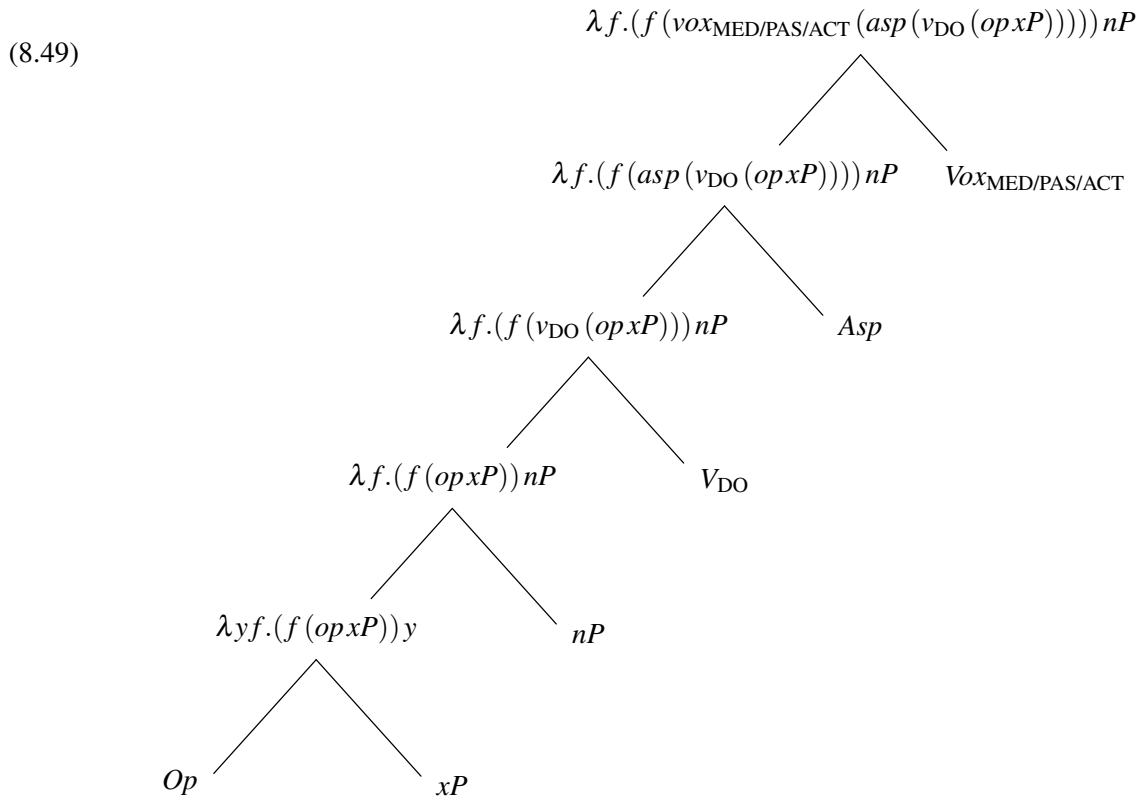
$$\langle +ac / +st, +in, +ec / ec, -ic / ic \rangle,$$

pertencentes a constituintes de estrutura

$$\lambda f. (f (Vox_{MED/PAS/ACT} (asp (v_{DO} (op xP))))) nP.$$

No constituinte acima, a voz ativa é caracterizada por uma configuração de traços especial e bem-definida. Entretanto, em que consiste a diferença entre as vozes média e ativa, se são licenciadas por estruturas semanticamente similares? A diferença, caso as nossas análise estejam corretas, localiza-se no operador asp , o único lugar em que pode haver variação semântica para o licenciamento das vozes, considerando-se as semânticas restritas de v_{DO} e dos demais operadores presentes na estrutura.

³⁵ ‘Maria tem um livro’ e ‘João deu um livro a Maria’ são exemplos de alternâncias de transitividade derivadas de $Prep_{HAVE}$.



Derivação das vozes média, passiva e ativa alternante

Proponho que: (a) a voz média é licenciada pelo operador *asp*, no caso de este assumir um valor ativo; e que (b) a voz passiva é licenciada por *asp* no caso de este assumir um valor estativo (como um particípio). O licenciamento de V_{MID} e de V_{PAS} , com base na atividade ou estatividade de *asp*, é compatível com a estrutura (a meu ver inevitável) dos traços semânticos e com a semântica das respectivas vozes.

Para deixar a teoria completa, é preciso fazer uma adição às condições de licenciamento da voz ativa: além das configurações estruturais e de traços até aqui analisadas é preciso acrescentar que a voz ativa é licenciada por um valor ativo de *asp*.

O licenciamento da voz passiva por *asp* dotado de valor estativo implica que não há ligação necessária entre a voz passiva e a sua expressão por meio de um particípio. No latim, por exemplo, a passiva se expressa por flexões verbais, de modo sintético, independente de particípio. O uso do particípio para a formação da voz passiva, como no português e no inglês, apenas reflete um recurso para se buscar um valor estativo para *asp*, dentro das possibilidades oferecidas pelo idioma.

Passaremos agora à análise das estruturas geradas pelas vozes média e passiva, relacionando-as com as estruturas ativas estudadas nos capítulos anteriores.

O modelo para a derivação das sentenças médias e passivas é analisado na seção 8.8.8, p. 146, em que derivamos um caso particular da voz média. Tomaremos o procedimento de

derivação como já conhecido.

- (a) As sentenças média ('Maria caiu-se') e passiva ('Maria foi caída') são agramaticais em ambas as vozes, devido ao traço $-ec$ da raiz inacusativa.
- (b) Nas estruturas alternantes, ativas e estativas, produzem-se as sentenças 'O rádio quebrou-se'/'O rádio está quebrado' e 'O café esfriou-se'/'O café foi esfriado'. As sentenças médias podem ser pouco usuais em PB, mas são gramaticais e não causam estranheza em 'O rádio quebrou-se sozinho' etc.. Os traços ec e ic das estruturas alternantes proporcionam a possibilidade de derivação das sentenças médias e também das ativas ('O rádio quebrou', 'O café esfriou'), com aproximadamente o mesmo significado. A escolha de uma ou de outra modalidade de expressão, no uso quotidiano da língua, é uma questão de hábitos linguísticos historicamente condicionados e não de correção gramatical.
- (c) As preposições P_{ALOC} e P_{COND} , ambas dotadas dos traços semânticos

$$\langle +ac, +in, +ec, -ic \rangle,$$

dão origem às estruturas denominais verbais locativa e condicional. A configuração de traços licencia de imediato as sentenças médias e passivas: 'O macaquinho engavetou-se'/'O macaquinho foi engavetado'; 'A conversa apimentou-se'/'A conversa foi apimentada'. Não é possível derivar uma sentença ativa dotada de sujeito interno, devido ao traço $-ic$. Há uma questão pendente há algum tempo: por que 'O livro engavetou-se' é semanticamente estranha, ainda que sintaticamente perfeita? A resposta está ligada ao papel temático do argumento externo da preposição (a posição de 'O macaquinho') neste tipo de estrutura.³⁶ Analisaremos a questão logo a seguir.

- (d) A preposição P_{LOC} (de traços semânticos $\langle +st, +in, +ec, -ic \rangle$) é base para a geração de 'João deu um livro a Maria'. Neste caso a derivação da passiva é imediata: 'Um livro foi dado a Maria'. Agora temos uma questão análoga à do parágrafo anterior: por que 'Um livro deu-se a Maria', a sentença média, é semanticamente anômala? A questão somente pode ser abordada ao se analisar o sujeito das estruturas verbais biargumentais.

8.9.3 O sujeito dos verbos biargumentais

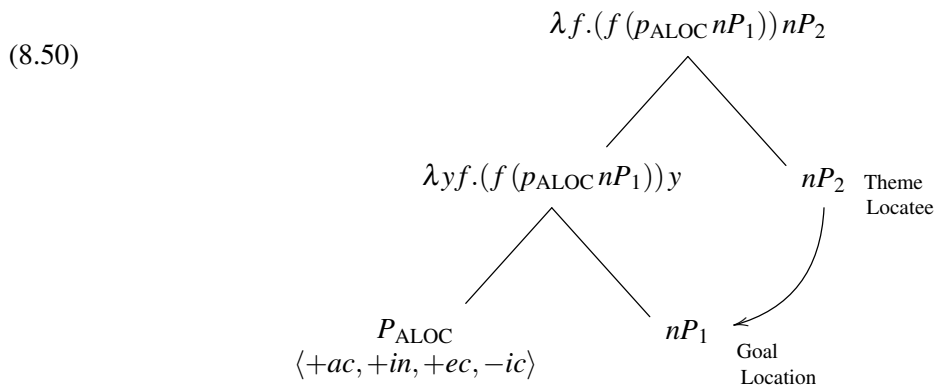
A análise dos sujeitos das estruturas verbais biargumentais lança luz sobre a interpretação das vozes média e passiva.

Primeiramente caracterizaremos a semântica dos sujeitos internos dos verbos denominais derivados das preposições P_{LOC} , P_{ALOC} e P_{COND} . O traço $-ic$, comum a todas essas preposições, faz com que somente possa haver sujeitos internos derivados das vozes média e passiva.

³⁶Os papeis temáticos são fundamentais à minha proposta. Recordar que os papeis temáticos são atribuídos por configurações estruturais e não por grades temáticas.

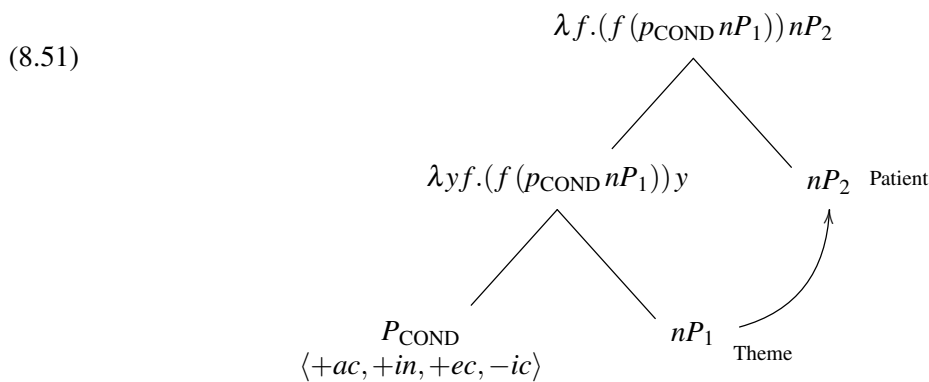
A semântica de P_{ALOC} , exibida no próximo diagrama, mostra nP_2 ocupando a posição estrutural de Theme (Tema), como origem de uma iniciativa/ato que afeta nP_1 , o ocupante da posição estruturalmente caracterizada como Goa (Meta). Semanticamente nP_2 deve ser dotado de agentividade, como iniciador de uma eventualidade.

A posição estrutural de nP_2 é a de sujeito das vozes média e passiva, explicando-se assim a agentividade esperada desses constituintes na posição de sujeito. Sentenças como 'Um livro engavetou-se', derivada de P_{ALOC} , violam a expectativa semântica relacionada ao sujeito.



Estruturalmente, P_{LOC} é semelhante a P_{ALOC} , a única diferença entre elas sendo a estatividade da primeira (traço $+st$, em lugar de $+ac$ da segunda). P_{LOC} deriva o verbo *dar* de duplo-complemento, tendo-se assim, de imediato, a passiva 'Um livro foi dado a Maria' (em que 'Um livro' é sujeito interno e 'Maria' ocupa a posição semântica de Goal). Verifica-se que a sentença média, 'Um livro deu-se a Maria' é semanticamente inesperada, por violar a expectativa de agentividade do sujeito.

Situação diversa mostra P_{COND} , de estrutura representada no próximo diagrama.



O argumento externo de P_{COND} , nP_2 , ocupa a posição semântica de Patient (eventualmente também a de Goal). De nP_2 nada se exige quanto à agentividade, que pode ou não existir. Assim, 'A conversa apimentou-se' e 'A conversa foi apimentada' não provocam estranheza semântica;

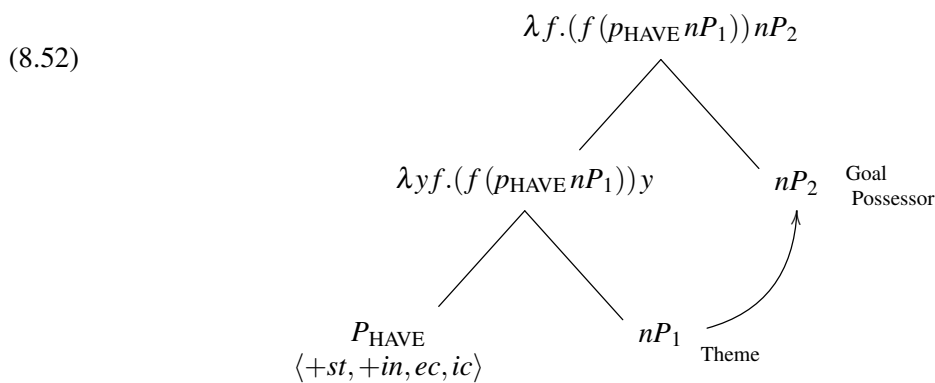
tampouco a provocam ‘O macaquinho apimentou-se’ e ‘O macaquinho foi apimentado’.

8.9.4 A semântica da voz média

A preposição P_{HAVE} , por ser dotada de duas semânticas claramente diferenciadas (verbo *dar* de duplo-objeto e verbo *ter*) possibilita a observação da semântica fundamental da voz média.

Como se observa do diagrama a seguir, P_{HAVE} possui argumento externo semanticamente caracterizado como Goal (Meta). Assim, não se espera que nP_2 seja dotado de traços de agentividade.

O traço *ic* licencia a interpretação de nP_2 como sujeito ativo, derivando o verbo *ter*, segundo o visto na seção 8.8.4, p. 141. Têm-se então sentenças como ‘Maria tem um livro’ e ‘A estante tem um livro’, em que a agentividade do sujeito (interno) não é estruturalmente relevante.



Como é previsto por nosso modelo, não há formação de sentenças médias ou passivas derivadas de *ter* (de *ter* derivado de P_{HAVE}), dado o seu licenciamento por *ic*.

Analisemos, entretanto, a estrutura de duplo-objeto, que não se realiza em português, ao menos em forma transitiva como a do inglês (‘John gave Mary a Book’).

Ocorre que a sentença ‘Maria deu-se um livro’ é gramatical em português e comum em muitos dialetos do PB (como o curitibano-italiano).³⁷ Portanto, (a) embora a variante de duplo-objeto de P_{HAVE} , licenciada por *ec* (verbo *dar* de duplo-objeto), inexista em português, (b) a variante média, dotada de sujeito interno, é normalmente derivada. O ponto fundamental, a ser ressaltado, é que a sentença média é derivada em português, mas somente o é com a semântica de *dar* (isto é, com a semântica “externa” de P_{HAVE}), justamente a semântica que não se expressa transitivamente em português.

O sujeito do verbo da voz média é interno, mas a semântica dessa voz é dada pelo verbo dotado de sujeito externo. Em termos mais precisos: a voz média licencia um sujeito interno,

³⁷Deste modo está sendo denominada a variedade de português geralmente falada pelas famílias de descendentes de italianos que colonizaram a região norte de Curitiba, estabelecendo-se nos bairros Mercês, Vista Alegre, Cascatinha, Santa Felicidade, São Braz e cercanias.

mas seleciona a expressão fonológica da estrutura derivada por *Voice*. Esta é uma condição mais forte do que o simples licenciamento pelo traço *ec*, que é agnóstico quanto a *Voice*.

Capítulo 9

O Princípio de Projeção Estendido

A binaridade estrita dos núcleos introduz sérias restrições às possibilidades de formular derivações e ao que se pode considerar uma estrutura linguística bem-formada. Em especial, os constituintes verbais, com a sua rica estrutura argumental, têm submetido o modelo a rigorosos testes, sobretudo levando-se em conta que as derivações devem assegurar a adequada interpretação temática dos argumentos em bases puramente estruturais.

Até o momento, princípios linguísticos independentemente motivados, implementados por meio de núcleos, asseguraram a estrutura estritamente binária das derivações, sem o recurso a qualquer estipulação. Um desses princípios é formalizado pelo núcleo *Voice*, que implementa a independência entre um argumento externo (futuro sujeito) e um núcleo verbal. *Voice* proporciona o argumento que torna sintaticamente bem-formado um constituinte construído sobre uma raiz inergativa.

Entretanto, é necessário considerar aquele que deve ser o limite inferior do modelo,¹ e talvez o limite inferior de qualquer teoria gerativa da linguagem, a derivação de expressões como

(15) Chove; *it rains*

(16) Existem/há bruxas; *there are witches*

Os exemplos mostram estruturas extensivamente analisadas na teoria linguística: sentenças em que expletivos ocupam a posição de sujeitos. Segundo a tradição gerativista, os sujeitos dessas sentenças foram inseridos por aplicação do *Princípio de Projeção Estendido* (EPP – Extended Projection Principle), que postula a obrigatoriedade de sentenças terem sujeito.

A motivação inicial de Chomsky ao formular o EPP, Chomsky (1981), foi a necessidade estrutural de determinados núcleos funcionais terem especificador; em especial, que algum elemento ocupe a posição de Spec, TP, o que resulta na obrigatoriedade de as sentenças terem sujeito. Em Chomsky (1995), o EPP foi reformulado em termos de uma teoria de verificação de

¹O limite superior deve situar-se na resolução de fenômenos de ligação que extrapolam os objetivos do trabalho.

traços, mas em Chomsky (2001) voltou-se à interpretação inicial, da obrigatoriedade de especificador para determinados núcleos funcionais. Extrapola os objetivos do trabalho fazer uma análise da teoria do EPP e remetemos os leitores interessados à literatura dedicada ao assunto.

De todo modo, as sentenças que a teoria gerativa (de vertente chomskiana) soluciona por aplicação do EPP coincidem com as estruturas que desafiam a natureza estritamente binária dos núcleos da presente proposta. É necessário, então, analisar a questão nos termos deste trabalho.

Assumirei que as sentenças anteriores, em português e em inglês, expressam os mesmos fenômenos sintáticos universais, em que pese a diferença de idioma e alguns pequenos detalhes superficiais (os sujeitos de ‘Chove’ e de ‘Existem bruxas’ não são completamente similares aos dos exemplos correspondentes em inglês). Procuraremos, então, determinar as estruturas universais dos exemplos, conjugando análises das sentenças em português e em inglês, partindo da ideia de que ambas as línguas proporcionam *insights* da estrutura em análise. Deste modo, tomaremos a raiz *chov* como subjacente a ‘Chove’ e ‘It rains’.

9.1 EPP e processos

Em primeiro lugar, analisemos a derivação de ‘Chove; *it rains*’, começando pela sentença em português.

Assumirei que a raiz *chov* designa um *processo espontâneo, destituído de causalidade interna ou externa*, e não-dotado de *incidência*. Trata-se, portanto, por hipótese, de uma raiz portadora dos traços $\langle +ac, -in, -ec, -ic \rangle$. Uma tal raiz será designada *raiz de atividade pura*, típica de verbos que expressam fenômenos meteorológicos.

Deste modo, “no caso geral”, em consequência do traço $-in$, a raiz *chov* é incompatível com o núcleo *Comp*, agregador de complementos. Fica então, de imediato, excluída a possibilidade de a raiz dar origem a estruturas dotadas de sujeitos originados em complementos. Ademais, o traço $-ec$ torna a raiz incompatível com *Voice*, e, assim, o argumento externo não pode provir deste núcleo. Ou seja, uma raiz de atividade pura não pode ser diretamente associada a argumentos, fato que a torna incompatível com sujeitos internos ou externos. Ademais, uma tal raiz é incompatível com V_{DO} , em consequência dos traços $-ec$ e $-ic$, e com V_{BE} , devido ao traço $+ac$.

Assim, a questão a ser investigada é a de como uma tal raiz, possuidora de traços que a tornam incompatível com núcleos verbalizadores (os traços $-in$, $-ec$ e $-ic$) pode dar origem a estruturas verbais dotadas de sujeito sintaticamente explícito, como em ‘It rains’, e também investigar a possibilidade de existência de um sujeito morfologicamente marcado na forma verbal, como em ‘Chove’, em que o sujeito se expressa pela pessoa do verbo.

Considerados os traços semânticos, uma raiz de atividade pura não pode servir de base à derivação de sentenças que tenham sujeitos sintaticamente gerados *por meio dos núcleos analisados até o momento*. É, então, necessário postular um núcleo especial para a finali-

dade: devido ao traço $+ac$ da raiz, o núcleo em análise deve ser compatível com a verbalização ativa (V_{DO}) e também capaz de criar a posição argumental necessária ao sujeito sentencial. Considerando-se que, neste caso, a origem do sujeito é devida ao próprio núcleo (isto é, o sujeito tem origem interna ao núcleo), o núcleo é dotado dos traços semânticos $\langle +ac, +in, -ec, +ic \rangle$: $+ac$, por compatibilidade com V_{DO} ; $+in$, $+ic$ e $-ec$, pela causalidade necessariamente interna. Designarei este núcleo por H_{AC} e a sua ação encontra-se ilustrada no diagrama a seguir. Este núcleo, no meu modelo, é parcialmente equivalente ao EPP da gramática gerativa; a equivalência é parcial, uma vez que somente se aplica às raízes de atividade pura.

(9.1)

$$\begin{array}{c}
 \lambda f.(f(h_{AC}(\lambda f.(f(r(ve\langle \sigma_i \rangle)))\langle \varphi_i \rangle)))yP \\
 \swarrow \quad \searrow \\
 H_{AC}rP \quad yP \\
 \swarrow \quad \searrow \\
 H_{AC} \quad \lambda f.(f(r(ve\langle \sigma_i \rangle)))\langle \varphi_i \rangle \\
 \langle +ac, +in, -ec, +ic \rangle \quad \langle +ac, -in, -ec, -ic \rangle
 \end{array}$$

O núcleo H_{AC} introduz, como seu argumento externo, o constituinte yP , a ser interpretado como sujeito (gramaticalmente gerado) por núcleos a serem introduzidos na derivação. Tendo-se em conta que o constituinte derivado

$$\lambda f.(f(h_{AC}(\lambda f.(f(r(ve\langle \sigma_i \rangle)))\langle \varphi_i \rangle)))yP$$

apresenta um constituinte raiz maximal,

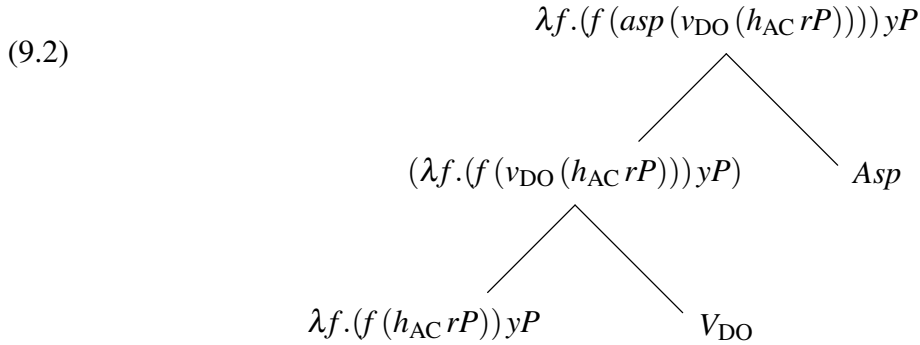
$$rP = \lambda f.(f(r(ve\langle \sigma_i \rangle)))\langle \varphi_i \rangle,$$

como argumento interno de h_{AC} , o constituinte derivado pode ser esquematicamente representado por

$$\lambda f.(f(h_{AC}rP))yP,$$

em que os traços semânticos $\langle \sigma_i \rangle$ e os traços fonológicos $\langle \varphi_i \rangle$ encontram-se subentendidos. Esta última representação será utilizada nas análises a seguir por apresentar maior simplicidade visual.

No constituinte $\lambda f.(f(h_{AC}rP))yP$, o operador h_{AC} estabelece a relação entre a raiz rP e o argumento gramatical yP ; os traços semânticos de h_{AC} , $\langle +ac, +in, -ec, +ic \rangle$, tornam o constituinte compatível com a verbalização ativa, mostrada no diagrama abaixo.



A estrutura gerada apresenta o constituinte yP como argumento externo da cadeia de operadores $f \circ asp \circ v_{DO} \circ h_{AC}$; tal cadeia de operadores (na qual f receberá valor(es) em etapas posteriores da derivação) interpreta yP como sujeito sentencial.

Em inglês, yP recebe *da cadeia de operadores* o valor fonológico *it*, puramente gramatical, observável em sentenças como ‘It rains’; em português, yP recebe, também gramaticalmente, valor equivalente à terceira pessoa,² valor que será expresso na forma do verbo, evidenciado em expressões como ‘Chove’. No primeiro caso, yP se manifesta por meio de um morfema livre (*it*); no segundo, por meio de um morfema ligado (desinência de terceira pessoa), dois modos de um constituinte se expressar.

É importante ressaltar que a cadeia $f \circ asp \circ v_{DO} \circ h_{AC}$ interpreta como sujeito a *qualquer* constituinte situado na posição yP , o seu argumento externo. Os traços $+ac$, $+in$, $-ec$, $+ic$, de H_{AC} , dotam o constituinte verbal derivado,

$$\lambda f.(f(asp(v_{DO}(h_{AC}rP))))yP,$$

de semântica inacusativa (no exemplo em português, trata-se de um constituinte derivado da raiz *chov*):³ (a) pode-se assim derivar sentenças como ‘Chove[m] canivetes’, em que a concordância peculiar e a posição pós-verbal do sujeito são habituais a verbos inacusativos em PB; (b) sujeitos agentivos, como em ‘João chove’, são comumente recusados em decorrência do traço $-ec$. No caso de H_{AC} se manifestar por meio de flexão verbal, a semântica inacusativa pode disparar um processo de concordância do verbo com o eventual sujeito interno, como ocorre em português. A semântica de número de H_{AC} é portanto inteiramente gramatical: o seu valor *default* (singular) é passível de se modificar por processos de concordância determinados pelo sistema da língua.

O constituinte yP , ao ser introduzido na derivação, é destituído de qualquer conteúdo

²O número recebe o valor *default* de singular, podendo entretanto alterar-se por concordância, como veremos algumas linhas adiante.

³Não se tem ainda efetivamente algo como *chover*, verbo dependente de outras fases da derivação.

lexical, semântico ou fonológico; sua estrutura interna reduz-se a

$$\text{null}P = \lambda f.(f(id\phi))\phi,$$

em que id é o operador identidade e ϕ é a sequência vazia. Este constituinte é computacionalmente possível, nada impede a sua construção, não sendo, assim, necessário estipular a sua existência computacional.

Ou seja: para determinação de um valor gramatical para yP , é necessário apenas estipular o comportamento gramatical de H_{AC} . Este resultado é metodologicamente importante: a introdução do núcleo H_{AC} equivale a uma única estipulação (o comportamento gramatical de H_{AC}), e não a duas (também o comportamento de yP).

9.2 EPP e estatividade

Analisemos agora

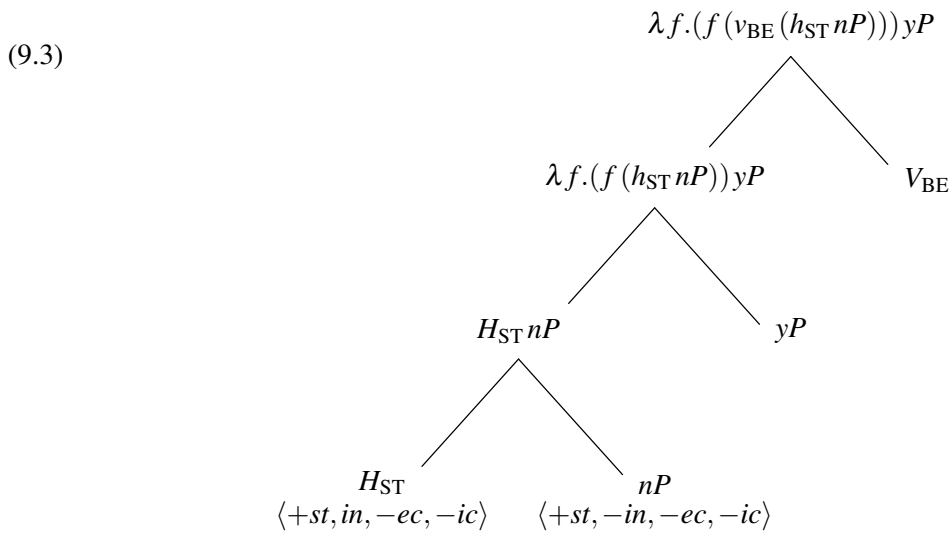
(17) Existem/há bruxas; *there are witches*

Do mesmo modo que nas análises das raízes de atividade pura, assumirei que as sentenças acima, independente do idioma, expressam a mesma estrutura universal, a afirmação de existência. Claramente, os verbos de existência aplicam-se a nPs . Uma vez que os únicos verbalizadores disponíveis são V_{DO} e V_{BE} , pode-se tomar como assegurado que os verbos de afirmação de existência expressam estruturas que envolvem V_{BE} , hipótese que adotaremos.

Assim, a derivação de 'Existem bruxas' inclui a participação de um nP (no exemplo, 'bruxas'), juntamente com o verbalizador V_{BE} . Há duas possibilidades diretas de derivação da sentença, e ambas podem ser descartadas de modo imediato. A aplicação direta $V_{BE}nP$ falha por não satisfazer a binaridade do núcleo, por falta de um segundo argumento. Também falha o processo inverso, em que o nP se aplica ao verbalizador (a aplicação nPV_{BE}): o argumento externo de V_{BE} é uma sequência fonológica, e não um nP , o que impossibilita a geração da fonologia verbal, levando a derivação a um estado de erro.

Deste modo, a geração das sentenças existenciais exige a colaboração de V_{BE} com um outro núcleo, a que designaremos provisoriamente por H . O núcleo H deverá participar da geração de sentenças como 'There are witches', em que 'There' ocupa a posição de sujeito sentencial e 'witches', a de complemento do núcleo. Tem-se, assim, a configuração $Hwitchesthere$, em que 'there' é um argumento gramatical. Portanto, H é um núcleo introdutor de argumentos gramaticais em estruturas existenciais; ou seja, é um núcleo de projeção argumental estativo, e o denotaremos por H_{ST} , o segundo núcleo de projeção do modelo proposto (o primeiro é H_{AC}). A estipulação do núcleo H_{ST} parcialmente equivale ao EPP da gramática gerativa, especificamente para as construções estativas.

Abaixo, temos a representação do núcleo H_{ST} , já incluído no ambiente verbal estativo,

V_{BE} .

O núcleo H_{ST} possui: (a) como argumento interno um constituinte nP dotado de conteúdo lexical; e (b) como argumento externo um constituinte yP , disponível para conteúdo puramente gramatical, ou, como veremos, também para constituintes lexicais. As propriedades dos argumentos nP e yP serão analisadas algumas linhas a seguir.

A estrutura acima deriva 'There are witches', em inglês: (a) yP recebe a interpretação gramatical de *there*, como argumento externo da cadeia de operadores $f \circ v_{BE} \circ h_{ST}$; e (b) a cadeia $f \circ v_{BE} \circ (h_{ST} witches)$ realiza fonologicamente v_{BE} como 'are' e o nP como 'witches'. Analogamente deriva-se, em português, 'Existem/há bruxas'. (A diferença entre *haver* e *existir* deve-se a outros núcleos situados acima de V_{BE} , possivelmente em colaboração com o núcleo *Asp*, não mostrado no diagrama.).

O diagrama anterior também evidencia que *existir* é estruturalmente diferente de *ser*. As diferenças estruturais entre esses verbos são analisadas em Benveniste (1950b), sob outro ponto de vista.

Passemos, agora, à análise do argumento interno de H_{ST} . O fato de H_{ST} aplicar-se a um argumento interno nP (no exemplo, bruxas) é estruturalmente fundamental. Se classificarmos os constituintes lexicais como *estativos* ou *ativos*, um nP representa uma forma da *estatividade pura*: possui estatividade, traço $\langle +st \rangle$, mas é destituído de incidência e causalidade, traços $\langle -in, -ec, -ic \rangle$. Ao considerar os constituintes nominais como estativos, adotamos visão análoga à de Hale & Keyser (2002).

De acordo com a seção 5.10, p. 58, V_{BE} é compatível apenas com constituintes dotados dos traços $\langle +st, in, -ec, -ic \rangle$. Estes, portanto, devem ser os traços semânticos do núcleo H_{ST} , para proporcionar a compatibilidade com V_{BE} . Isto é: o núcleo H_{ST} cria o ambiente apropriado para a derivação da afirmação de existência, em que um constituinte nominal maximal é verbalizado por V_{BE} ; o efeito semântico introduzido por H_{ST} é o da afirmação de existência.

9.3 Modelo proposto e EPP

No modelo computacional proposto, os sujeitos sentenciais são licenciados por relações de compatibilidade existentes entre os núcleos verbalizadores, V_{DO} e V_{BE} , e os traços semânticos ac , st , in , ec e ic , *exceto* nos casos de raízes de atividade pura (aquelas dotadas dos traços $\langle +ac, -in, -ec, -ic \rangle$) e de estatividade pura (as possuidoras dos traços $\langle +st, -in, -ec, -c \rangle$). Nestes casos, ao se afirmar a existência de processos ('Chove'/'It rains') ou de estados ('Existem bruxas'/'There are witches'), há a necessidade de postulação de núcleos relacionadores especiais, H_{AC} e H_{ST} ; as combinações de traços semânticos mostram-se insuficientes tanto para a verbalização ativa quanto para a estativa.

O modelo computacional especifica correlações sistemáticas entre as configurações de traços semânticos dos constituintes e as possibilidades de derivação sintática e faz previsões sobre as configurações de traços em que é impossível o licenciamento estrutural de sujeitos e a verbalização direta de constituintes, como é o caso dos constituintes portadores dos traços de atividade e de estatividade puros.

Ou seja, o modelo computacional desenvolvido neste trabalho necessita ser complementado com (supostos) primitivos funcionais da GU (H_{AC} e H_{ST}) exatamente nos mesmos pontos em que as derivações da gramática gerativa somente podem prosseguir mediante a postulação de um princípio especial, o EPP. Considero tal fato como uma evidência positiva de correção da presente proposta.

Parte V

Conclusão

Capítulo 10

Considerações finais

A presente pesquisa tem por objetivo buscar um paradigma mínimo de representação sintática. Para tanto parte do pressuposto de que o fenômeno linguístico é composto de processos. Este é o motivo da escolha do cálculo-lambda como paradigma de representação: o caráter fundamental do cálculo-lambda é o de especificação abstrata e mínima de processos.

Alternativamente, a teoria dos conjuntos, por sua grande economia conceitual, é também candidata a fundamentar a representação linguística. Há entretanto um fato a considerar: a teoria dos conjuntos é uma teoria de “coisas” (objetos).

Cada uma dessas teorias expressa aspectos complementares da realidade: aspectos dinâmicos (cálculo-lambda, ao representar processos) e aspectos estáticos (teoria dos conjuntos, ao representar objetos).

Embora a teoria dos conjuntos, ao longo da história da Gramática Gerativa, tenha recebido a preferência como fundamento formal da representação linguística, creio que a continuidade das pesquisas mostrará a maior adequação do cálculo-lambda para a finalidade.¹

A teoria desenvolvida neste trabalho possui três conceitos primitivos: um primitivo computacional (o *pair*), um primitivo sintático (o núcleo) e um primitivo lexical (o elemento vocabular). Todas as derivações sintáticas compõem-se exclusivamente desses elementos.

O caráter fortemente derivacional da teoria apresenta muitas implicações cujos aspectos mais gerais sumariamos a seguir.

Primeiramente, é importante observar que não existe algo como o “Princípio da Projeção” (PP), como em (Chomsky 1981, p. 29). Os elementos lexicais, e após eles as raízes,

¹As discussões a respeito do cálculo-lambda e da teoria dos conjuntos como fundamento de teorias formais é muito mais ampla do que sugerem as breves observações dos parágrafos anteriores. Inicialmente o cálculo-lambda foi desenvolvido por Church como parte de um projeto para fundamentação da matemática, em uma pesquisa de doutorado orientada por D. Hilbert, o mais proeminente matemático do final do Sec. XIX e início do Sec. XX. Buscava-se então uma alternativa à teoria dos conjuntos e à lógica como fundamento da matemática, a partir da ideia de “processo efetivo” (algoritmo composto de um número finito de passos). Após os resultados de Gödel, a abordagem de Church mostrou-se insuficiente para a formulação de teorias matemáticas completas e isentas de contradições, como também assim mostraram-se a teoria dos conjuntos e a lógica. Entretanto, o cálculo-lambda, ainda que não tenha atingido as expectativas iniciais do seu autor, veio a tornar-se um dos principais recursos da Ciência da Computação e é o fundamento de várias linguagens de programação, como Lisp, ML, Haskell, Erlang e outras.

não possuem propriedades de subcategorização de itens lexicais. Por princípio, nenhum constituinte é dotado de qualquer especificação de ambientes em que pode ser inserido: um constituinte pode sinalizar uma (in)compatibilidade de associação a outros itens, mas não pode portar critérios gramaticais que determinem a especificação desses itens.

Some-se a isto outro fato fundamental: não há níveis sintáticos (LF, *Logical Form*; DS, *Deep Structure*; SS, *Surface Structure*) no modelo proposto. Existe um único nível sintático, que não é apenas um nível representacional, uma vez que um constituinte, além de ser uma representação, é também uma computação, completada até o ponto em que se encontra e em estado suspenso para operações futuras. Ou seja, o nível de representação sintática é também um nível de execução computacional. Existe uma única “realidade” derivacional, o constituinte situado no topo da derivação. A história derivacional de um constituinte encontra-se codificada em sua estrutura, a qual contém toda a informação fonológica e semântica do constituinte. Potencialmente, o constituinte possui uma expressão fonológica e uma expressão semântica (algo como a LF), mas por meio de processos pós-sintáticos.

Portanto, não há, na proposta, os elementos fundamentais presentes no princípio da projeção: níveis de representação e propriedades de subcategorização lexical.

Outro fato fundamental é o de os constituintes serem destituídos de informações temáticas, embora sejam dotados de traços semânticos. Toda a informação temática decorre das propriedades estruturais do constituinte. A noção de grade temática é alheia ao modelo proposto.

Por limitação de tempo, este texto apenas relata tópicos relacionados à estrutura argumental; alguns outros temas foram deixados à parte para continuação futura.

Referências Bibliográficas

- Arad, M. (2002), 'Locality constraints on the interpretation of roots: the case of Hebrew denominal verbs', *Ms.* p. 8. 70, 71
- Baker, M. C. (1985), *Incorporation: A theory of grammatical function changing*, MIT, Massachusetts Institute of Technology. 24, 41
- Baker, M. C. (1996), *The polysynthesis parameter*, Oxford University Press, New York. 62
- Baker, M. C. (2004), *Lexical categories: verbs, nouns, and adjectives*, Cambridge. 47, 62
- Barendregt, H. P. (1984), *The lambda calculus: its syntax and semantics*, North-Holland, Amsterdam. 9
- Barendregt, H. P. (1990), *Functional programming and lambda calculus*, Vol. B: Formal models and semantics, Elsevier, Amsterdam, chapter 7, pp. 321–363. 9, 41
- Barendregt, H. P. (1997), 'The impact of the lambda calculus in logic and computer science', *The Bulletin of Symbolic Logic* **3**(2), 181–215. 3
- Benveniste, Émile. (1950a), *Ativo e médio no verbo*, in Benveniste (1991), chapter 14, pp. 183–191. 100, 156
- Benveniste, Émile. (1950b), *A frase nominal*, in Benveniste (1991), chapter 13, pp. 163–182. 169
- Benveniste, Émile. (1991), *Problemas de Linguística Geral I*, Pontes Editores, São Paulo. 174
- Burzio, L. (1986), *Italian Syntax: A Government-Binding approach*, Dordrecht, Kluwer. 97
- Carnielli, W. & Epstein, R. (2006), *Computabilidade, funções computáveis, lógica e os fundamentos da Matemática*, Editora UNESP, São Paulo. 8
- Chomsky, N. (1955/1975), *The logical structure of linguistic theory*, Plenum Press, New York. 127, 128, 139
- Chomsky, N. (1965a), *Aspects of the theory of syntax*, MIT Press, Cambridge. 97
- Chomsky, N. (1965b), *The formal nature of language*, Cambridge University Press, chapter 5, pp. 102–142. 32
- Chomsky, N. (1970), *Remarks on nominalization*, Braisdell, Waltham, Mass., pp. 184–221. 4
- Chomsky, N. (1981), *Lectures on government and binding: the Pisa lectures*, Mouton de Gruyter, Berlin. 4, 5, 21, 24, 28, 32, 130, 164, 172

- Chomsky, N. (1993a), *A minimalist program for linguistic theory*, in Chomsky (1995), pp. 167–217. 3, 17
- Chomsky, N. (1993b), *A minimalist program for linguistic theory*, in Hale & Keyser (Edts.) (1993), pp. 1–52. 28
- Chomsky, N. (1995), *The minimalist program*, MIT Press, Cambridge. 4, 21, 30, 68, 130, 164, 175
- Chomsky, N. (2001), *Derivation by phase*, MIT Press, Cambridge, Mass., chapter 1, pp. 1–52. 165
- Church, A. (1941), *The calculi of lambda conversion*, Princeton University Press. 3
- Dowty, D. (1979), *Word meaning and the Montague Grammar*, D. Reidel Publishing Comp., Dordrecht, The Netherlands. 25
- Eco, U. (1991a), *Semiótica e filosofia da linguagem*, Ática, São Paulo. 75
- Eco, U. (1991b), *Tratado geral de semiótica*, Perspectiva, São Paulo. 18
- Fillmore, C. J. (1968), *Em favor do caso*, Rio de Janeiro, pp. 275–365. 5, 23, 24
- Givón, T. (1984), *Syntax: A functional-typological introduction*, John Benjamins, Amsterdam, The Netherlands. 24
- Gruber, J. S. (1965), *Studies in lexical relations*, Doctoral thesis, MIT. 23
- Hale, K. & Keyser (Edts.), S. J. (1993), *The VIEW from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, MIT Press, Cambridge, Mass. 28, 175
- Hale, K. & Keyser, S. J. (1993), *On argument structure and the lexical expression of syntactic relations*, in Hale & Keyser (Edts.) (1993), pp. 53–110. 4, 28, 32
- Hale, K. & Keyser, S. J. (1998), *The basic elements of argument structure*, in Hale & Keyser (2002), chapter 1, pp. 1–27. 5, 32
- Hale, K. & Keyser, S. J. (2001), *Aspect and the syntax of argument structure*, Oxford University Press, pp. 11–41. 28
- Hale, K. & Keyser, S. J. (2002), *Prolegomenon to a theory of argument structure*, The MIT Press, Massachusetts. 5, 28, 30, 31, 41, 92, 102, 106, 118, 125, 130, 131, 132, 133, 144, 169, 175
- Halle, M. & Marantz, A. (1993), *Distributed morphology and the pieces of inflection*, in Hale & Keyser (Edts.) (1993), pp. 111–176. 4
- Halle, M. & Marantz, A. (1994), *Some key features of distributed morphology*, Papers on Phonology and Morphology, MIT Working Papers in Linguistics 21, 275–288. 6, 68
- Hankin, C. (2004), *An introduction to lambda calculi for computer scientists*, King's College Publications, Imperial College London. 41
- Harley, H. (1998), 'You're having me on!: aspects of *have*'. 137

- Harley, H. (2002a), 'A minimal(ish) linking theory', *University of Arizona Maryland Minimalist Mayfest*(Handout). 116, 117, 123
- Harley, H. (2002b), 'Possession and the double object construction', *University of Arizona Linguistic Variation Yearbook*(2), 29–68. 22, 23, 134, 135, 137, 142, 143
- Harley, H. (2007), *The bipartite structure of verbs cross-linguistically (or: Why Mary can't exhibit John her paintings)*, ABRALIN, UFMG, Belo Horizonte. 129, 134, 137, 142, 143
- Harley, H. & Noyer, R. (April 1999), 'State-of-the-article: Distributed morphology', *Glott International* 4(4), 3–9. 4, 32
- Jackendoff, R. (1972), *Semantic interpretation in generative grammar*, MIT Press, Cambridge, Massachussetts. 24
- Kratzer, A. (1996), *Severing the external argument from its verb*, Kluwer, Indiana University, pp. 109–137. 99, 106
- Larson, R. K. (1988), 'On the double object construction', *Linguistic Inquiry* 19(3), 335–391. 129, 130
- Levin, B. & Rappaport Hovav, M. (2005), *Argument realization*, Cambridge University Press, Cambridge. 5, 22, 23, 25, 26
- Marantz, A. (1996), 'Cat' as a phrasal idiom: consequences of late insertion in Distributed Morphology', *Cambridge, Mass.; MIT Press* pp. 1–24. 4
- Marantz, A. (1998), 'No escape from syntax: don't try morphological analysis in the privacy of your own lexicon', *Proceedings of the 1998 Penn Linguistics Colloquium, ed. by Alexis Dimitriadis, available from Penn Working Papers in Linguistics* pp. 201–205. 4
- Marantz, A. P. (1981), *On the nature of grammatical relations*, Massachusetts Institute of Technology, MIT. 130
- Partee, B. H., ter Meulen, A. & Wall, R. E. (1990), *Mathematical methods in linguistics*, Kluwer Academic Publishers, The Netherlands. 9
- Peirce, C. S. (1867), 'On a new list of categories', *Proceedings of the American Academy of Arts and Sciences* (7), 287–298. 3
- Perlmutter, D. M. (1978), 'Impersonal passives and the unaccusative hypothesis', *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society* (BLS 4), 157–189. 97
- Pesetsky, D. (1995), *Zero syntax: experiencers and cascades*, MIT Press, Cambridge, Massachusetts. 26
- Santaella, L. (2000), *Teoria geral dos signos – como as linguagens significam as coisas*, Editora Pioneira, São Paulo. 3
- Saussure, F. (1978), *Curso de Linguística Geral*, Publicações Dom Quixote, Lisboa. 68
- Seki, L. (2000), *Gramática do Kamaiurá: Língua Tupi-Guarani do Alto Xingu*, Editora da Unicamp, Campinas. 57

- Stowell, T. A. (1981), *Origins of phrase structure*, MIT, Doctoral thesis. 5, 23
- Turing, A. M. (1937), 'Computability and λ -definability', *The Journal of Symbolic Logic* (2), 153–163. 8
- Van Valin, R. (1993), *A synopsis of role and reference grammar*, John Benjamins, In *Advances in Role and Reference Grammar*, ed. R. Van Vallin, Amsterdam, The Netherlands. 24
- Vendler, Z. (1967), *Linguistics in philosophy*, Cornell University Press, Ithaca, New York. 25

Índice Remissivo

- ||, 13
 - ||, veja Concatenação
 - definição, 13
- Late insertion*, 102
- Pattern*, 70
- Conflation*, 30, 116, 131
- Pôr*, 147
- Pattern*, 69, 73
 - e elemento vocabular, 73
- Shell*
 - Larsonian*, 129
- κ , 59
- Asp*, 91
- P_{HAVE}
 - papeis temáticos, 140
 - traços semânticos, 140
- P_{LOC} , 138, 160
 - traços semânticos, 138
- Voice*, 107
- Vox_{ACT} , 102, 107
- Vox*, 66
- σ_i , 59
- φ_i , 59
- w_k , 59
 - assinatura funcional, 59
- b*
 - assinatura funcional, 59
- e*, 60
- i*, 60
- k*, 59
 - assinatura funcional, 59
- p*, 59
- s*, 59
- H_{AC} , 166, 170
- H_{St} , 170
- N_{URA} , 48
- P_{ALOC} , 161
- P_{COND} , 161
 - semântica, 149
 - traços semânticos, 149
- P_{GEN} , 151
 - traços semânticos, 152
- P_{HAVE} , 140, 162
- P_{LOC} , 161
- Vox_{ACT} , 101, 115, 148, 155
 - licenciamento, 155
 - semântica fundamental, 155
- Vox_{MID} , 146, 156–158
- Vox_{PAS} , 157, 158
 - traços semânticos, 157
- V_{BE} , 151, 170
 - derivação estativa, 89
 - e predicado nominal, 95
 - e raiz de est. básica, 95
 - parametrização, 58
 - traços de licenciamento, 95
 - traços licenciadores, 96
 - traços semânticos, 151
- V_{DO} , 170
 - e raiz estativa, 95
 - traços licenciadores, 96
- w_k , 61
 - assinatura funcional, 61
- Asp*, 120
- N*
 - e formação lexical, 73
- R*
 - e formação lexical, 73
- Ve*
 - e formação lexical, 73
- Voice*, 115, 116
- Vox*, 100, 101, 116, 120, 153
 - interpretação verbal, 100
- V*
 - e formação lexical, 73
 - parametrização, 58
- ac*, 47
- ec*, 47
- first*
 - utilizada em *infix*, 14
- ic*, 47
- infix*
 - definição, 14
 - exemplo, 14
- in*, 47
- pair*, 41

- prefix*, 11
 - definição, 14
 - exemplo de derivação, 14
- second*
 - utilizada em *infix*, 14
- st*, 47
- suffix*
 - definição, 13
 - exemplo de derivação, 13
- Ve*, 68
 - definição, 77
 - parametrização, 80
- β , 80
 - traços semânticos de *R*, 80
- o, 12
 - see Composição, 12
- c*
 - ordem de constituintes, 63
- Abstração, 9
- Adjetivo, iii, 31, 47, 69, 76, 88, 89
 - e *pattern*, 73
 - estrutura adjetival lexical, 152
 - nome complexo, 94
- Adposição, iii, 47, 136
- Aplicação, 9
- Argumento, 12
 - de expressão-lambda, 12
- Argumento externo
 - introd. estrutural, 129
- Cálculo-lambda, iii, 3, 8
 - composição de operadores, 12
 - computação abstrata, 8
 - e fundamentos da matemática, 172
 - e processos, 172
 - elementos, 9
 - fenômeno linguístico, 8
 - não-tipado, 16
 - ordem de parâmetros, 39
 - tipado, 16
- Cadeia
 - de núcleos, 31, 130
 - de operadores, 92
- Caso
 - sistema nominativo-acusativo, 96
- Categoria gramatical, 58
- Categoria lexical, 47
- Composição, 12
 - de funções, 12
 - de operadores, 12
- Concatenação, 13
- Concha
 - larsoniana, 129
- Conflation, 30, 31, 117
 - e *Merge*, 30
 - proposta de Hale e Keyser, 102
- Constituinte, iii, 4, 44
 - assoc. de forma e signif., 21
 - componente fonológico, 19
 - componente semântico, 19
 - composição de núcleos, 51
 - definição, 44
 - e configs. sintáticas, 6
 - e subsistemas da GU, 21
 - estrutura, 52
 - estrutura de nome, 20
 - expr. fonológica, 51
 - história derivacional, 49, 85
 - lexicalização, 81
 - maximal, 45, 78
 - repr. de alto nível, 49
 - repr. de baixo nível, 49
 - semântica lexical, 6
 - significado enciclopédico, 75
- Contexto, 22, 36, 41
 - externo, 52
 - propr. estruturais, 60
- Dativo, 128
 - repr. por adjunto, 128
- Diagrama, 57
- Elemento vocabular, 18, 67, 68
 - caráter estativo, 76
 - conexão entre forma e signif., 68
 - definição, 77
 - estr. do constituinte, 74
 - evidência do hebraico, 70
 - evidência empírica, 69
 - exemplo de derivação, 76
- EPP, 166, 170
 - Gramática Universal, 170
- Estrutura
 - de duplo-complemento, 129, 136
 - de duplo-objeto, 129, 136
 - genitiva, 152
- Estrutura inacusativa
 - traços semânticos, 97
- Estrutura inergativa
 - traços semânticos, 105
- Estrutura verbal
 - estr. v. fundamentais, 2
- Genitivo

- estrutura gen., 152
 - preposição, 151
- Gramática Universal, 17, 47
 - e EPP, 170
 - e operadores, 47
 - e sintaxe, 21
 - e traços sem., 47
 - traços sem. universais, 47
- h, 59
 - h, *veja* Núcleo
- Infixação, 13
- Inserção tardia, 102
- Léxico, 49
- Lexicalização, 49, 81, 84
- Máquina de Turing, 8
- N
 - estrutura, 83
- Núcleo, iii, 46
 - assinatura fonológica, 131
 - assinatura funcional, 59
 - assinatura-p, 131
 - categorizador, 66
 - de formação lexical, 74
 - definição, 59
 - e cat. gramatical, 58
 - e constituinte, iii
 - e diagrama, 57
 - e traços sem. e fon., 47
 - efeito introduzido, 73
 - especificação de parâmetros, 59
 - exemplo de núcleo parametrizado, 48
 - expressão contextual, 51
 - fonologia defectiva, 131
 - método de análise, 151
 - modalizador, 66
 - operador, 4
 - ordem de pronúncia de args., 63
 - parametrização, 12
 - parametrizado, 46
 - parametrizado, exemplo, 48
 - relacionador, 66
 - result. bem-sucedido, 51
 - resultado, 56
 - resultado lambda, 60
 - tipos de, 66
- Nome, iii, 47
 - constituente complexo, 94
 - derivação, 83, 86
 - derivado de adjetivo, 85
 - derivado de raiz estativa., 85
 - entidade complexa, 21
 - estrutura, 83
 - parametr. de semântica e fonologia, 86
- Nominalizador
 - estrutura interna, 55
- Operações morfológicas, 13
 - infixação, 13
 - prefixação, 13
 - sufixação, 13
- Operador
 - composto, 21
 - de alto nível, 58
 - e contexto externo, 56
 - sinonímia, 13
- Operador característico, 51
 - e contexto externo, 53
 - es, 55
 - estrutura interna, 52, 56
 - exemp. estrutura interna, 53
 - introdução de *w*, 55
 - visibilidade sintática, 52
- Parâmetro, 12, 46
 - de expressão-lambda, 12
 - exterior, 46, 59
 - interior, 46, 59
- Paradigma gramatical, 55, 56
 - e argumento externo, 56
 - e contexto externo, 56
 - ordem de pronúncia de args., 63
- Parametrização, 12
 - de expressão-lambda, 12
- Particípio
 - adjetivo verbal, 100
- Português-Brasileiro, 3
 - dialeto-padrão, 3
- Prefixação, 11, 13
- Preposição, 136
 - e arg. externo, 62
- Processamento postergado, 102
- R, 80
- Raiz, 69, 70
 - de estatividade pura, 87
 - de atividade pura, 165
 - de atividade pura , 170
 - de estatividade básica, 89, 118
 - de estatividade pura, 170
 - derivação, 80

- e *pattern*, 69
 - e incidência, 88
 - e verbos inacusativos, 95
 - estativa básica, 94
 - exemplo de derivação, 76
 - inacusativa, 81, 97
 - não-estativa, 95
 - nominal, 87
 - parametrização, 80
 - trilítera, 73
 - verbalização por V_{BE} , 94
- Semiótica, 2
- Significado enciclopédico, 75
- Sintaxe, 21
 - subsistema da GU, 21
- Sufixação, 13
- Sujeito
 - constituente verbal, 103
 - inacusativo, 103
 - posição estrutural, 92, 100, 103
- Traços
 - fonológicos, 59
 - semânticos, 47, 59
- Transitividade
 - alternância, 112
- Valência, 12
 - de expressão-lambda, 12
- Valor de argumento, 12
 - de expressão-lambda, 12
- Variável, 10
 - ligada, 10
 - livre, 10
- Verbo, iii, 5, 47
 - ativo alternante, 112
 - biargumental, 128, 136
 - de duplo-complemento, 127
 - de duplo-objeto, 127
 - de mudança de estado, 112
 - denominal, 127, 136
 - ditransitivo, 127
 - e *pattern*, 73
 - gramatical, 127
 - inacusativo
 - derivação, 100
 - incoativo, 112
 - inergativo, 105
 - lexical, 127
 - transitivo, 108
 - transitivo lexical, 5
 - transitivo puro, 108
- Voice, 111
- Vox, 100
 - categorização verbal, 154
 - média, 146
- Voz, 100
 - ativa, 100, 153
 - derivação, 159
 - licenciamento, 159
 - semântica fundamental, 100
 - média, 100, 138, 146, 153, 156
 - derivação, 159
 - licenciamento, 159
 - semântica, 162
 - semântica fundamental, 156
 - passiva, 100, 153, 156
 - derivação, 159
 - licenciamento, 159